



UNIVERSIDAD AUTÓNOMA DE MADRID

PROGRAMA DE DOCTORADO EN BIOCIENCIAS MOLECULARES

**CARACTERIZACIÓN MOLECULAR DE UNA COHORTE DE
258 FAMILIAS ESPAÑOLAS CON RETINOSIS
PIGMENTARIA AUTOSÓMICA DOMINANTE**

TESIS DOCTORAL

M^a Inmaculada Martín Mérida

Madrid, 2018



UNIVERSIDAD AUTÓNOMA DE MADRID

Facultad de Medicina

Departamento de Bioquímica

**CARACTERIZACIÓN MOLECULAR DE UNA COHORTE DE
258 FAMILIAS ESPAÑOLAS CON RETINOSIS
PIGMENTARIA AUTOSÓMICA DOMINANTE**

TESIS DOCTORAL

Memoria presentada por **M^a Inmaculada Martín Mérida**, licenciada en Bioquímica, para optar
al título de Doctora por la Universidad Autónoma de Madrid

Directoras de Tesis:

Dra. Carmen Ayuso García y Dra. Marta Cortón Pérez

Realizado en el Instituto de Investigación Sanitaria- Fundación Jiménez Díaz (IIS-FJD)

Resumen/Abstract

Las distrofias hereditarias de retina (DR), con una prevalencia de 1/3000, son un conjunto de enfermedades raras producidas por la degeneración de los fotorreceptores y representan una de las principales causas de ceguera en el mundo occidental. Las DR presentan una elevada heterogeneidad, tanto clínica como genética, con más de 260 genes asociados hasta el momento. La retinosis pigmentaria (RP) es la forma más frecuente de las DR y se puede presentar bajo todos los tipos de herencia descritos, en los que las formas dominantes (adRP) corresponden aproximadamente al 15% de las familias RP en España. El número de genes implicados, actualmente 28, la aparición tardía de los síntomas, la expresividad variable y la penetrancia incompleta hace que el diagnóstico molecular de la adRP sea complejo.

En esta Tesis Doctoral se describe el cribado genético más completo realizado hasta la fecha en familias con adRP de origen europeo, utilizando estrategias basadas en secuenciación masiva (NGS) y un estudio complementario de variaciones en el número de copia (CNVs). Gracias al uso de paneles de genes implicados en distintas formas de DR, en nuestra cohorte de 258 familias españolas, hemos obtenido una tasa global de caracterización genética del 60%. El rendimiento del diagnóstico genético en la adRP aumentó desde el 38% obtenido con técnicas tradicionales de cribado al 65% con el uso de la NGS, cuando ambas aproximaciones fueron utilizadas como primer abordaje de estudio.

Los genes más prevalentes fueron *RHO*, *PRPF31*, *RP1* y *PRPH2*, representando el 62% de las familias caracterizadas. La mayoría de las mutaciones encontradas fueron de tipo *missense*. Es de destacar que el estudio de CNVs, mediante MLPA y análisis de los datos de profundidad de lectura obtenidos mediante NGS, nos permitió caracterizar al 3% del total de las familias con CNVs principalmente en los genes *PRPF31* y *CRX*.

Hemos profundizado en el papel de *PRPF31*, el segundo gen más prevalente, sobre distintos aspectos de su etiopatogenia en el desarrollo de adRP. Nuestro trabajo confirmó que la penetrancia incompleta es un fenómeno muy frecuente en familias portadoras de mutaciones en *PRPF31*, afectando hasta el 65% de las familias. Además, el estudio en estas familias del minisatélite MSR1 adyacente al promotor de *PRPF31*, propuesto como principal modulador de la penetrancia en este gen, puso de manifiesto que debe de existir un control poligénico de la penetrancia del gen, que permita explicar los casos portadores de variantes en este gen que se mantienen asintomáticos.

En conclusión, la aplicación de forma sistemática de la NGS a una cohorte de familias con adRP ha permitido aumentar la tasa de caracterización en estas familias, la reclasificación de algunas de las mismas, así como la elaboración de un algoritmo más efectivo para su diagnóstico.

Inherited Retinal Dystrophies (IRD), with a prevalence of 1/3000, represent a group of rare diseases characterized by the degeneration of photoreceptors and are one of the main causes of blindness in the Western world. IDRs present a high clinical and genetic heterogeneity, with more than 260 associated genes described up to now. Retinitis pigmentosa (RP), the most frequent form of IDR, can present under any of the described modes of inheritance, where autosomal dominant forms (adRP) correspond to approximately 15% of RP families in Spain. The number of genes involved, currently 28, the late onset of symptoms, the variable expression and the incomplete penetrance makes molecular diagnosis of adRP complicated.

This Doctoral Thesis describes the most complete genetic screening carried out to date in families of European origin with adRP by using strategies based on next-generation sequencing (NGS) and the study of copy number variations (CNVs). In our cohort of 258 Spanish families, we obtained a global genetic characterization rate of 60%, by using panels of genes responsible for different forms of IDRs. When using each strategy as first study approach the genetic diagnosis yield in adRP increased from 38%, obtained with traditional screening techniques, to 65% with the use of NGS.

The most prevalent genes were *RHO*, *PRPF31*, *RP1*, and *PRPH2*, accounting for 62% of the characterized families. Most of the mutations found were missense variants. The study of CNVs, through MLPA and a read-depth analysis of the data obtained by NGS, allowed us to characterize 3% of the families, with CNVs mainly in the *PRPF31* and *CRX* genes.

We studied the role of *PRPF31*, the second most prevalent gene in our cohort, on different aspects of its etiopathogenesis in the adRP development. Our work confirmed that incomplete penetrance is a very frequent phenomenon in families with mutations in *PRFP31*, affecting up to 65% of families. In addition, the study of the minisatelite MSR1 adjacent to the promoter of *PRPF31* in these families, proposed as the main modulator of the penetrance in this gene, showed that there must be a polygenic control of the gene penetrance that allows to explain the cases carrying variants in this gene that remain asymptomatic.

In conclusion, the systematic application of the NGS to a cohort of families with adRP has allowed us not just for increasing the characterization rate of adRP families, but also for the reclassification of some of them, as well as for elaborating a more effective algorithm for adRP molecular diagnosis.

Índice

Resumen/Abstract	1
Índice de tablas y figuras	11
Abreviaturas	17
1. Introducción	21
1.1. Morfología y fisiología de la retina	23
1.1.1. Tipos celulares de la retina	24
1.1.2. Los fotorreceptores	25
1.2. La fototransducción y el ciclo visual	26
1.2.1. Cascada de la fototransducción.....	26
1.2.2. Ciclo visual	27
1.3. Las distrofias de retina	29
1.4. Retinosis pigmentaria	30
1.4.1. Aspectos clínicos.....	30
1.4.2. Aspectos genéticos	32
1.5. Retinosis pigmentaria autosómica dominante	33
1.6. El gen <i>PRPF31</i> y la penetrancia incompleta	37
1.7. Estudio molecular de la adRP	40
2. Objetivos.....	43
3. Material y métodos	47
3.1. Sujetos y muestras	49
3.2. Cribado mutacional de variantes de un solo nucleótido (SNV)	50
3.2.1. Panel customizado de genes de RP mediante HaloPlex.....	52
3.2.2. Panel de genes de RD mediante la tecnología MIPs	54
3.2.3. Secuenciación del exoma completo	56
3.2.4. Exoma clínico	56
3.2.5. Priorización y clasificación de las variantes.....	57

3.3. Secuenciación de Sanger	59
3.4. Análisis de CNVs.....	60
3.4.1. MLPA.....	60
3.4.2. Análisis de CNVs con datos de NGS	60
3.4.3. Arrays de hibridación genómica comparada.....	61
3.4.4. <i>Long-Range</i> PCR.....	62
3.5. Análisis de haplotipos	63
3.6. Análisis de expresión génica	64
3.7. Análisis de CNVs del minisatélite MSR1 adyacente al promotor de <i>PRPF31</i>	64
3.8. Bases de datos y herramientas bioinformáticas	65
4. Resultados.....	69
4.1. Cribado mutacional de 258 familias españolas con adRP mediante distintas tecnologías moleculares.....	72
4.2. Análisis de variaciones en el número de copia (CNV)	73
4.2.1. Análisis de los reordenamientos mediante MLPA.....	75
4.2.2. Análisis de los reordenamientos mediante datos de NGS	76
4.2.3. Validación de los reordenamientos en <i>PRPF31</i> y <i>CRX</i> mediante aCGH.....	76
4.3. Tasa global de detección de mutaciones en 258 familias con adRP en población española.	80
4.3.1. Espectro mutacional	81
4.3.2. Reclasificación clínica y genética de las familias con cambios recesivos o ligados al cromosoma X	84
4.4. Estudio del gen <i>PRPF31</i> en familias españolas con adRP	86
4.4.1. Espectro mutacional de <i>PRPF31</i> en nuestra población española.....	86
4.4.2. Dos reordenamientos en <i>PRPF31</i> segregando en la misma familia	90
4.4.3. Contribución del elemento repetido MSR1 a la penetrancia de la enfermedad de las mutaciones de <i>PRPF31</i>	91
4.4.4. Correlación genotipo-fenotipo para las familias con variantes en <i>PRPF31</i>	92

5. Discusión	97
5.1. Datos globales de la caracterización genética de la adRP.....	99
5.2. Algoritmo diagnóstico actual para las familias con adRP.....	108
5.3. Importancia del gen <i>PRPF31</i> en la adRP	109
6. Conclusiones	113
7. Bibliografía	117
8. Anexos	135
Anexo 1. Listado de SNVs y pequeñas in/dels probablemente patogénicas identificadas en nuestra cohorte.....	137
Anexo 2. Árboles y estudio de segregación de las familias reclasificadas	140
Anexo 3. Árboles y estudio de segregación de las familias adRP con mutación en el gen <i>PRPF31</i>	144
Anexo 4. Publicaciones derivadas de esta tesis doctoral	148

Índice de tablas y figuras

Índice de tablas

Tabla 1. Genes asociados a adRP	35
Tabla 2. Listado de genes usado asociados a distrofias de retina	51
Tabla 3. Condiciones de la <i>Long-Range</i> PCR para la amplificación de la duplicación en el gen <i>PRPF31</i>	62
Tabla 4. Microsatélites del gen <i>PRPF31</i>	63
Tabla 5. Microsatélites del gen <i>CRX</i>	64
Tabla 6. Condiciones de la PCR para la amplificación de <i>MSR1</i>	65
Tabla 7. SNV causales tras el reanálisis del panel de NGS usando la tecnología de HaloPlex....	73
Tabla 8. Grandes reordenamientos genómicos encontrados en nuestra cohorte.....	74
Tabla 9. Descripción del rendimiento diagnóstico en función de la estrategia molecular seguida	80
Tabla 10. Variantes patogénicas en 3 o más familias en nuestra población de 258 familias con adRP	82
Tabla 11. Prevalencia de cada uno de los genes identificados en nuestra cohorte de 258 familias españolas con adRP.	83
Tabla 12. Listado de SNVs probablemente patogénicas identificadas en nuestra cohorte en genes no asociados a adRP	85
Tabla 13. Mutaciones en el gen <i>PRPF31</i> encontradas en la cohorte de 258 familias españolas con adRP.....	88
Tabla 14. Características clínicas de 31 pacientes afectados de 18 familias caracterizadas con una mutación en el gen <i>PRPF31</i>	93
Tabla 15. Descripción del fenotipo en función de la edad de aparición de los síntomas, en pacientes con mutación en <i>PRPF31</i>	95

Tabla 16. Porcentaje de caracterización y prevalencia de los genes <i>RHO</i> , <i>PRPF31</i> , <i>RP1</i> y <i>PRPH2</i> en distintas poblaciones de adRP.....	104
--	-----

Índice de figuras

Figura 1. Anatomía del ojo humano.....	23
Figura 2. Estructura de la retina humana.	24
Figura 3. Esquema de los tipos de fotorreceptores de la retina.....	26
Figura 4. Esquema de la fototransducción y el ciclo visual.....	28
Figura 5. Clasificación de las distrofias de retina	29
Figura 6. Diagrama de genes asociados las distrofias de retina	30
Figura 7. Representación esquemática del proceso de <i>splicing</i>	37
Figura 8. Distribución continua de la expresión de <i>PRPF31</i>	38
Figura 9. Control de la expresión génica de <i>PRPF31</i>	39
Figura 10. Representación esquemática de la preparación de las librerías del método de captura basado en HaloPlex.....	53
Figura 11. <i>Pipeline</i> implementado en nuestro laboratorio.....	54
Figura 12. Representación esquemática del diseño y flujo de trabajo del método de captura basado en MIPs	55
Figura 13. Representación esquemática de la preparación de las librerías mediante el exoma clínico	57
Figura 14. Algoritmo llevado a cabo para la clasificación de las variantes	59
Figura 15. Imagen de un individuo heterocigoto 3/4 para el minisatélite MSR1	65
Figura 16. Algoritmo usado para estudiar a las 258 familias con adRP	72
Figura 17. Árbol y segregación de la familia RP-0932 con dos reordenamientos en <i>PRPF31</i>	75

Figura 18. Resultados de los reordenamientos mediante el análisis de la profundidad de lectura de los datos de NGS	77
Figura 19. Acotamiento de las deleciones del gen <i>PRPF31</i> mediante aCGH.	78
Figura 20. Acotamiento de las deleciones del gen <i>CRX</i> mediante aCGH.....	79
Figura 21. Estrategia molecular seguida para el cribado del gen <i>PRPF31</i> en una cohorte española de 258 familias con adRP	86
Figura 22. Análisis de haplotipos en dos familias con la misma variante <i>nonsense</i> en <i>PRPF31</i> . 89	
Figura 23. Análisis de haplotipos en dos familias con la misma variante de <i>splicing</i> en <i>PRPF31</i> .89	
Figura 24. Expresión del gen <i>PRPF31</i> en la familia RP-0932	90
Figura 25. Validación mediante PCR larga de la duplicación en tándem en la familia RP-0932. 91	
Figura 26. Densidad de secuencias tipo Alu y genes en los distintos cromosomas.....	102
Figura 27. Algoritmo diagnóstico propuesto para las familias con adRP.....	108

Abreviaturas

aCGH	<i>Comparative genomic hybridization array</i> . Array de hibridación genómica comparativa
ACL	Amaurosis congénita de Leber
ADN/DNA	Ácido desoxirribonucleico
adRP	Retinosis pigmentaria autosómica dominante
ARN/RNA	Ácido ribonucleico
ARNm/ mRNA	ARN mensajero
arRP	Retinosis pigmentaria autosómica recesiva
AV	Agudeza visual
AVC	Alteración de la visión cromática
CV	Campo visual
CN	Ceguera nocturna
CNV	<i>Copy number variation</i> . Variación en el número de copias
CSNB	<i>Congenital Stationary Night Blindness</i> . Ceguera nocturna congénita estable
DBC	Distrofia de bastones-conos
DCB	Distrofia de conos-bastones
DGGE	<i>Denaturing gradient gel electrophoresis</i> . Electroforesis de gel de gradiente desnaturalizante
DM	Distrofia macular
DMAE	Degeneración macular asociada a la edad
DR	Distrofias de retina
EPR	Epitelio pigmentario de la retina
ERG	Electrorretinograma
GMP	Guanosín monofosfato
GMPc	Guanosín monofosfato cíclico
GDP	Guanosín difosfato
GTP	Guanosín Trifosfato
INL	<i>Inner nuclear layer</i> . Capa nuclear interna
LCA	<i>Leber congenital amaurosis</i> . Amaurosis congénita de Leber
LOF	<i>Loss of Function</i> . Pérdida de función
LORD	<i>Late-onset retinal degeneration</i> . Degeneración retiniana de inicio tardío
MAF	<i>Minor allele frequency</i> Frecuencia alélica menor
MIPs	<i>Molecular inversion probes</i> Sondas moleculares de inversión
MSR1	<i>Minisatellite repeat element</i> Elemento minisatélite repetido

MLPA	<i>Multiplex ligation-dependent probe amplification</i> Sonda de amplificación múltiple dependiente de ligación
NGS	<i>Next generation sequencing</i> . Secuenciación masiva
NMD	<i>Non- sense mediated decay</i> . Decadencia mediada por mutaciones sin sentido
OCT	Tomografía de coherencia óptica
OMIM	<i>Online Mendelian Inheritance in Man</i> Herencia mendeliana humana online
ONL	<i>Outer nuclear layer</i> . Capa nuclear externa
PI	Penetrancia incompleta
PRPF31	<i>Pre-mRNA processing factor 31</i>
qPCR	<i>quantitative PCR</i> . PCR cuantitativa
RP	Retinosis pigmentaria
SINE	<i>Short interspersed element</i> . Elemento corto interpuesto
snRNAs	<i>Small nuclear RNAs</i> . ARN nuclear pequeño
SNP	<i>Single-nucleotide polymorphisms</i> Polimorfismo de único nucleótido
SNV	<i>Single nucleotide variant</i> . Variación de único nucleótido
sRP	Retinosis pigmentaria esporádica
SSCP	<i>Single-strand conformation polymorphism</i> . Polimorfismo de conformación de hebra única
STR	<i>Short tandem repeat</i> . Repetición corta en tandem
VUS	<i>Variant of uncertain significance</i> (variantes de significado incierto)
VREF	Vitreorretinopatía exudativa familiar
WES	<i>Whole exome sequencing</i> . Secuenciación de Exoma completo
WGS	<i>Whole genome sequencing</i> . Secuenciación de Genoma completo
xLRP	Retinosis pigmentaria ligada al cromosoma X

Introducción

1.1. Morfología y fisiología de la retina

Los órganos de los sentidos nos permiten recibir estímulos del medio que nos rodea, y los transforman en estímulos nerviosos que son transmitidos al cerebro. De los cinco sentidos, la visión es considerado el más importante. La percepción visual es llevada a cabo por la retina, tejido que recubre la superficie interna del ojo, que se encuentra constituida por células neuronales, entre ellas los fotorreceptores, y células de soporte (Figura 1). Cuando la luz incide en la retina se produce una cascada de eventos bioquímicos y enzimáticos, conocida como fototransducción, que finalmente desencadena un impulso nervioso que se transmite por el nervio óptico hasta la corteza visual en el cerebro (den Hollander *et al.*, 2010; Lledo-Riquelme *et al.*, 2010).

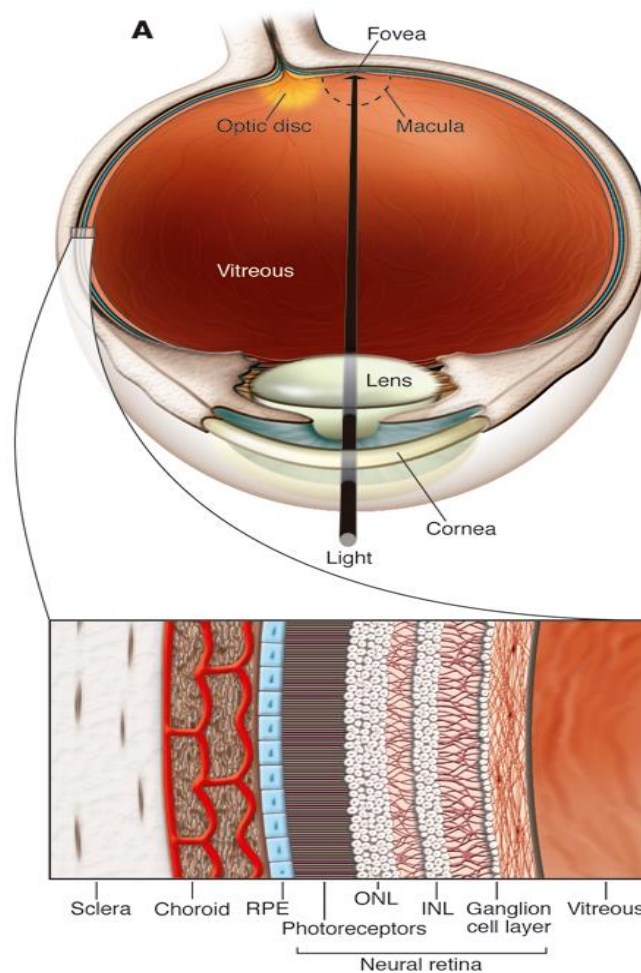


Figura 1. Anatomía del ojo humano. Detalle de las distintas capas concéntrica a nivel del polo posterior, incluyendo la estructura de la retina. RPE: epitelio pigmentario de la retina. INL (*inner nuclear layer*): capa nuclear interna. ONL (*outer nuclear layer*): capa nuclear externa. Imagen adaptada de den Hollander *et al.*, 2010.

1.1.1. Tipos celulares de la retina

La retina está formada por una porción neurosensorial y por el epitelio pigmentario retiniano (EPR). Está constituida por tres capas que contienen los cuerpos neuronales y dos capas, denominadas plexiformes, que contienen las interacciones sinápticas (Figura 2). La capa nuclear externa contiene los cuerpos celulares de los fotorreceptores, que son células neuronales modificadas sensibles a la luz. La capa nuclear interna contiene los cuerpos celulares de las células horizontales, bipolares, amacrinas e interplexiformes. La capa de células ganglionares contiene los cuerpos celulares de estas células (Hartong *et al.*, 2006; Lledo-Riquelme *et al.*, 2010).

El EPR está firmemente unido a la capa vascular subyacente llamada coroides y está en íntimo contacto con los fotorreceptores. Participa en el ciclo visual, proceso continuo en el que los retinoides usados en la fototransducción se reciclan, así como en el aporte de nutrientes a los fotorreceptores (den Hollander *et al.*, 2010).

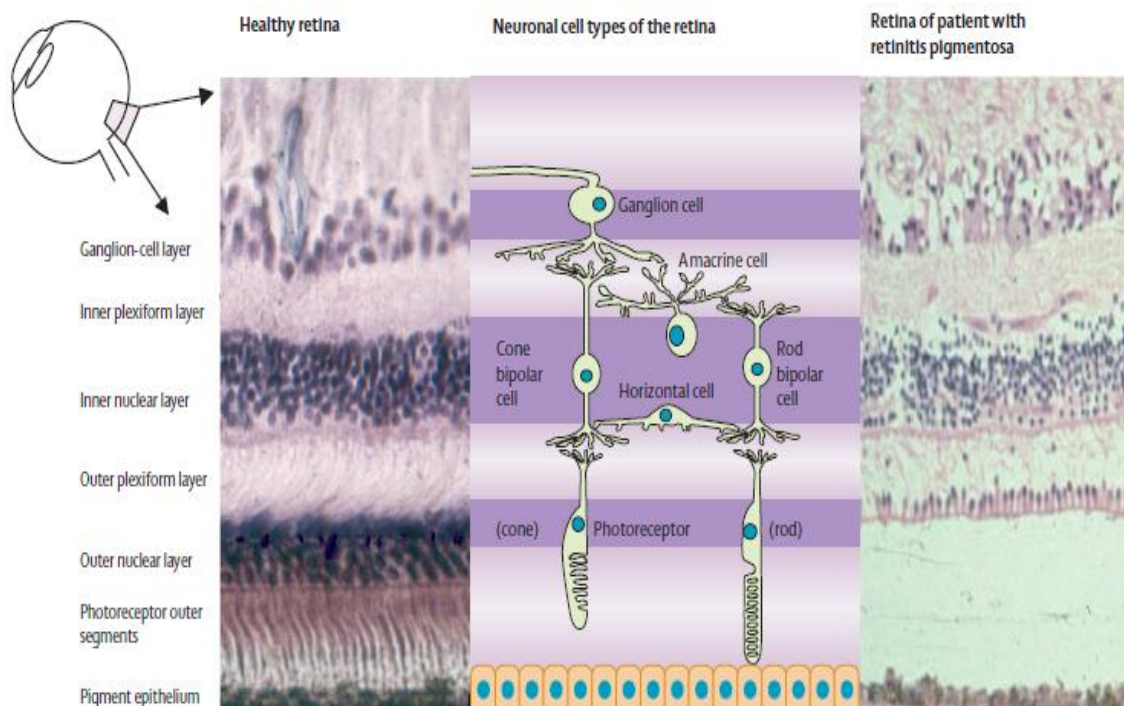


Figura 2. Estructura de la retina humana. Sección de retina humana formada por diferentes capas en una retina normal y un paciente con retinitis pigmentaria; y representación de los distintos tipos celulares. Imagen adaptada de Hartong *et al.*, 2006.

1.1.2. Los fotorreceptores

En la retina existen dos tipos de células fotosensibles o fotorreceptores, que presentan diferencias estructurales, funcionales y de localización:

- Los **bastones** que median la visión en blanco y negro y la visión nocturna o en penumbra. Presentan un único tipo de pigmento visual denominado rodopsina.
- Los **conos** son responsables de la visión en color y la agudeza visual durante la luz diurna. Existen tres tipos de conos en función del tipo de pigmento (opsina) que presenten: rojos, verdes y azules, que absorben a diferentes longitudes de onda y confieren la capacidad de visión cromática (den Hollander *et al.*, 2010).

En la retina humana existen alrededor de 6 millones de conos y unos 120 millones de bastones por lo tanto, existe una proporción de 20 bastones por cada cono. Aunque los bastones y conos se distribuyen por toda la retina, los bastones se localizan en la periferia de la retina mientras que los conos se concentran en la región central de la retina, llamada mácula, donde la máxima densidad de conos se encuentra en su parte más central o fovea (Lledo-Riquelme *et al.*, 2010).

La estructura de los fotorreceptores consta de un **segmento externo** donde se localiza el pigmento visual y por ello, donde se produce la fototransducción (Figura 3). Está constituido por discos membranosos formados, en el caso de los conos, por invaginaciones de la membrana plasmática, densamente empaquetadas llamadas lamelas, y en el caso de los bastones envueltos por la membrana plasmática, llamados discos. El **segmento interno** de los fotorreceptores contiene los orgánulos celulares típicos (aparato de Golgi, retículo endoplasmático, mitocondrias, etc.), está conectado con el segmento externo por un **cilio conector**. Además, se distingue el **cuerpo celular** donde se encuentra el núcleo, un **axón** y el **terminal axónico** (llamado esférula en los bastones y pedículo en los conos), que contiene gran cantidad de vesículas sinápticas.

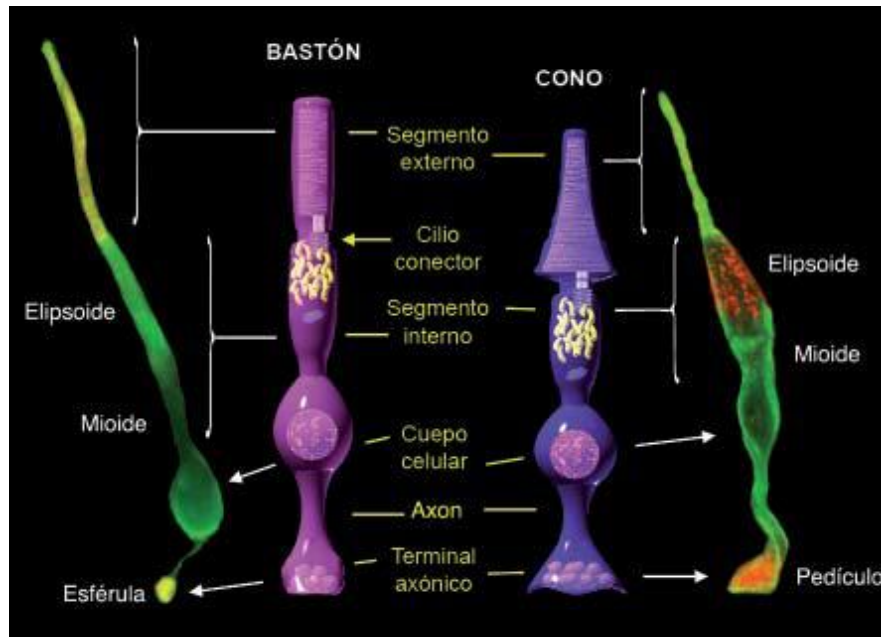


Figura 3. Esquema de los tipos de fotorreceptores de la retina (Lledo-Riquelme *et al.*, 2010).

1.2. La fototransducción y el ciclo visual

Cuando la luz incide en el ojo, atraviesa todas las capas de la retina desde las células ganglionares hasta alcanzar el segmento externo de los fotorreceptores. La señal luminosa captada por los fotorreceptores es procesada y transmitida a las células bipolares, siendo modulada por las células horizontales. Las células bipolares transmiten la información a las células amacrinas y ganglionares, los axones de este último tipo celular constituyen el nervio óptico, a través del cual se envía la información visual al cerebro (Hartong *et al.*, 2006; Lledo-Riquelme *et al.*, 2010).

1.2.1. Cascada de la fototransducción

La fototransducción es un proceso mediante el cual un fotón de luz genera una respuesta nerviosa en los fotorreceptores (Arshavsky *et al.*, 2002), gracias a la activación de los pigmentos visuales en las membranas de los discos del segmento externo, que activan a su vez una cascada de reacciones enzimáticas que serán transformadas en un impulso nervioso.

En los bastones, la rodopsina está formada por una proteína transmembrana llamada opsina, unida covalentemente a un cromóforo derivado de la vitamina A llamado retinal, que en fase de oscuridad se encuentra en la forma *cis* (Figura 4). En fase luminosa, mediante la mediación de un fotón de luz, el retinal isomeriza a la forma *trans*, lo que causa un cambio conformacional de

la rodopsina a su estado activo y cataliza el intercambio de los nucleótidos GDP por GTP de una proteína G denominada transducina, que consta de tres subunidades (α , β y γ). La subunidad α de la transducina con GTP activará a la enzima fosfodiesterasa (*PDE6A/PDE6B*, compuesta por cuatro subunidades: α , β y 2γ), que estimula la degradación de GMPc, ocasionando el cierre de los canales de Na^+ localizados en la membrana plasmática del fotorreceptor, quedando hiperpolarizada debido al aumento de Na^+ extracelular. Este cambio en el potencial de membrana produce el cierre de los canales de Ca^{2+} dependientes de voltaje, codificados por los genes *CNGA1/CNGB1*, lo que implica una disminución de entrada de Ca^{2+} en la sinapsis y una disminución de la secreción del neurotransmisor glutamato por parte de los fotorreceptores. En fase de oscuridad, la sinapsis con las células bipolares mediada por glutamato es inhibitoria, mientras que la llegada de un fotón de luz provoca el efecto contrario, estimulando a las células bipolares, e iniciándose un impulso nervioso que es transmitido primero a las células ganglionares y posteriormente a la corteza visual.

Después de cada cascada activada por la luz, se debe de producir una reacción de inactivación con el fin de devolver al fotorreceptor a su estado de reposo, esencial para que se puedan generar respuestas ante nuevos estímulos luminosos y cambios en los niveles de iluminación. La inactivación de la rodopsina se lleva a cabo por la rodopsina quinasa y, como consecuencia, la rodopsina inactiva se une a la arrestina, codificada por el gen *SAG*. La inactivación de la PDE se lleva a cabo por la hidrólisis de GTP unido a la subunidad α de la transducina y la unión del complejo multiproteico RGS9-1.G β 5.R9AP. Por otro lado, los niveles citoplasmáticos de GMPc en oscuridad se regeneran a partir de GTP gracias a la actividad de la enzima guanilato ciclasa, codificada por el gen *GUCY2D*. Esta enzima está estimulada por las proteínas activadoras de la guanilato ciclasa, GCAP, que detectan el descenso intracelular de Ca^{2+} secundario al cierre de los canales con la luz (Mendes *et al.*, 2005; Lledo-Riquelme *et al.*, 2010).

1.2.2. Ciclo visual

Los pigmentos visuales inactivos también deben de ser regenerados para mantener el correcto funcionamiento de los fotorreceptores. La inactivación de la rodopsina mediante su fosforilación y unión a arrestina provoca su descomposición generándose opsina y *todo-trans*-retinal. Esta última molécula es transportada al EPR a través de la proteína transportadora ligante de ATP, codificada por el gen *ABCA4*, para convertirse en *11-cis*-retinal, tras varias reacciones enzimáticas llevadas a cabo por las proteínas RDH12, LRAT, RGR, RPE65 y RDH5 en el espacio citoplasmático del segmento externo del fotorreceptor y el EPR. El *11-cis*-retinal vuelve a los

bastones para unirse a la opsina libre y regenerar el pigmento visual, disponible para la llegada de un nuevo fotón (den Hollander *et al.*, 2010; Lledo-Riquelme *et al.*, 2010) (Figura 4).

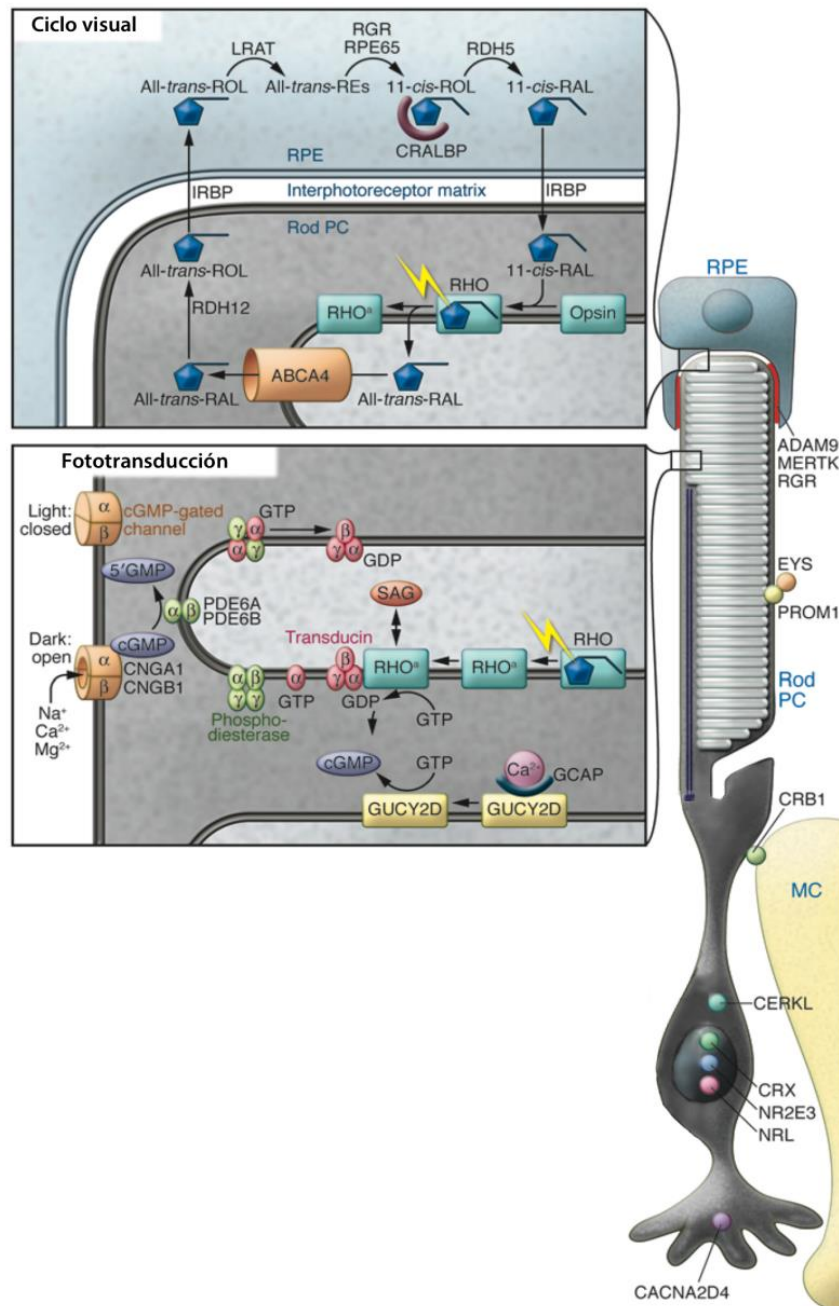


Figura 4. Esquema de la fototransducción y el ciclo visual. Imagen adaptada de Hollander *et al.*, 2010.

1.3. Las distrofias de retina

Las distrofias hereditarias de retina (DR) son un grupo de enfermedades genéticas que están caracterizadas por la pérdida generalmente progresiva de los fotorreceptores (Ayuso & Millan, 2010). Con una prevalencia de 1:3000, son consideradas enfermedades raras; sin embargo representan la mayor causa de ceguera en el mundo occidental en adultos en edad laboral (Liew *et al.*, 2014).

Las DR se suelen clasificar según el tipo celular que se vea afectado de manera primaria. Así, en las formas centrales, aparece una afectación predominante de la mácula debido a la degeneración de los conos mientras que, en las formas periféricas, se ven afectados los bastones, siendo la forma más frecuente la retinosis pigmentaria (RP) (den Hollander *et al.*, 2010) (Figura 5).

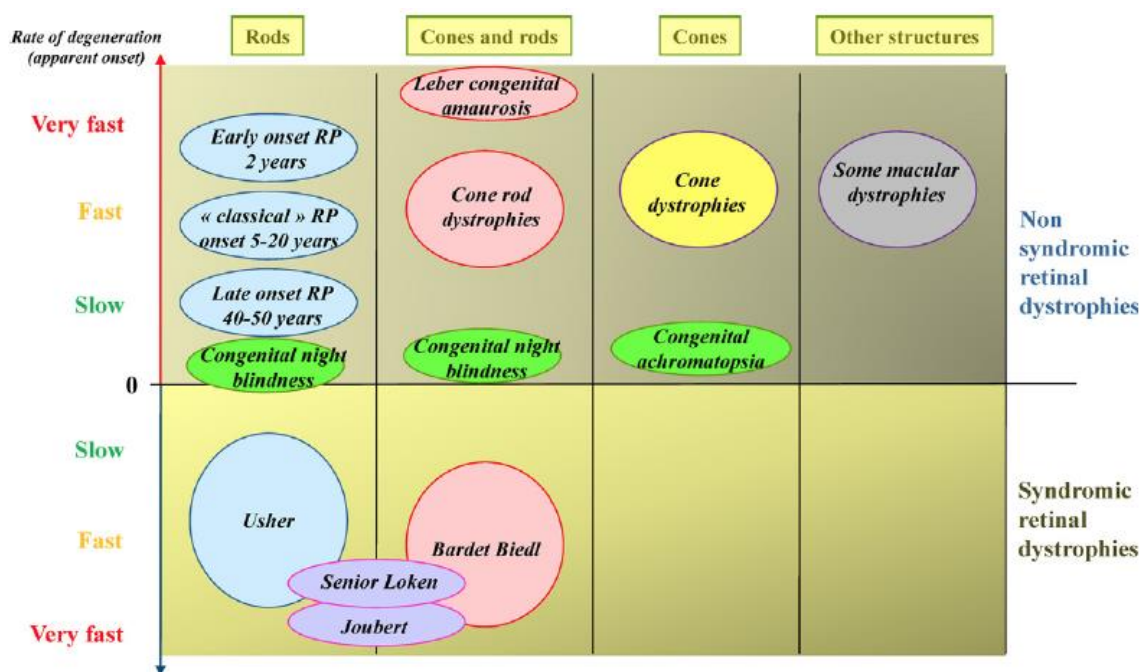


Figura 5. Clasificación de las distrofias de retina, en función del tipo de fotorreceptor afectado, grado de degeneración, comienzo de los síntomas, y condición sindrómica o no sindrómica (Hamel *et al.*, 2014).

En general, las DR presentan una elevada heterogeneidad tanto clínica como genética, donde distintas mutaciones en un mismo gen pueden dar lugar a distintos fenotipos y a su vez, el mismo fenotipo estar causados por mutaciones en distintos genes (Figura 6). Se han descrito más de 261 genes asociados a DR según la base de datos RetNet (<https://sph.uth.edu/retnet/sum->

dis.htm#B-diseases, enero 2018) que explican alrededor del 60% (Farrar *et al.*, 2017) de los casos, incluyendo más de 89 genes responsables de formas sindrómicas de DR. Por todo ello, el diagnóstico genético de este grupo de patologías resulta complejo.

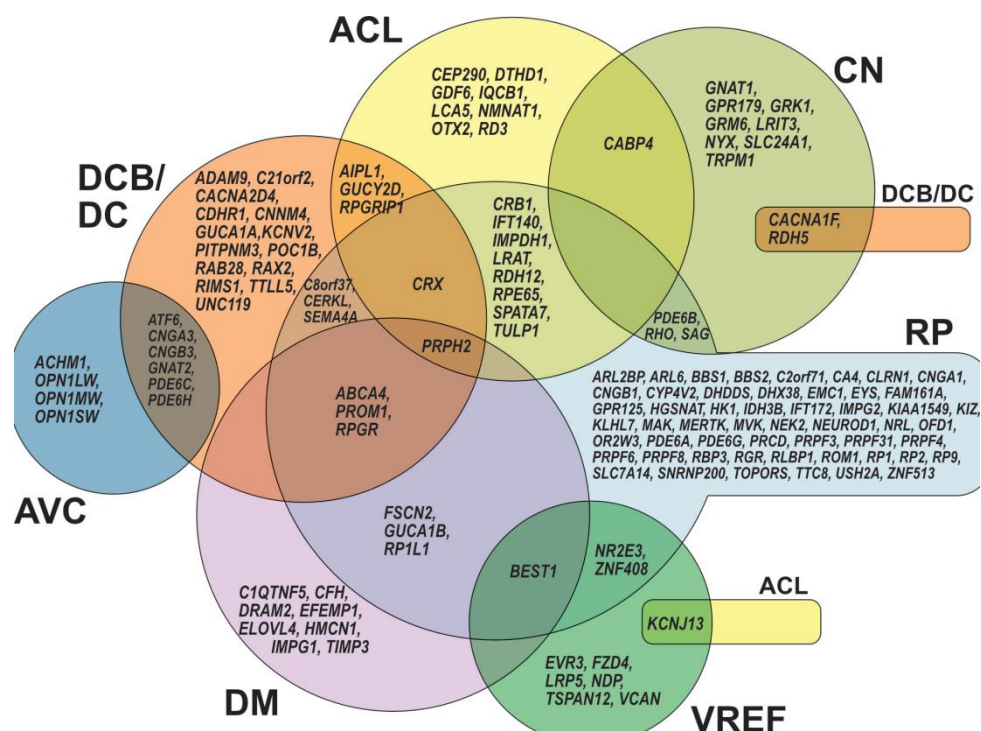


Figura 6. Diagrama de genes asociados a las distrofias de retina. Las zonas solapantes indican cómo un mismo gen puede dar lugar a diferentes fenotipos. RP: retinosis pigmentaria; CN: ceguera nocturna; ACL: amaurosis congénita de Leber; DCB/DC: distrofia de conos-bastones y distrofia de conos; AVC: alteración de la visión cromática; DM: degeneración macular; VREF: vitreooretinopatía exudativa familiar. Imagen adaptada de Berger *et al.*, 2010.

1.4. Retinosis pigmentaria

Con una prevalencia de 1 en 4000 (Hamel, 2006), la retinosis pigmentaria es la forma más común dentro de las DR, siendo responsable del 85 al 90% del total de los casos (Ayuso & Millan, 2010). La RP es conocida también como distrofia de bastones-conos, debido a que se produce una degeneración primaria de los bastones seguida de una degeneración secundaria de los conos, lo que puede conducir a una ceguera total (Ali *et al.*, 2017).

1.4.1. Aspectos clínicos

La RP es una enfermedad muy variable entre individuos, incluso dentro de la misma familia, con respecto tanto a la edad de inicio de los síntomas como a las características y los hallazgos oftalmológicos (Hartong *et al.*, 2006; Hamel *et al.*, 2014; Guía de Práctica Clínica para las

Distrofias Hereditarias de Retina, 2017). Mientras algunos pacientes presentan sintomatología de forma precoz desde la infancia, otros permanecen asintomáticos hasta la mitad de la edad adulta (Figura 5).

Los síntomas visuales que presentan los pacientes son indicativos de la pérdida gradual de los dos tipos de fotorreceptores y suelen aparecer de forma bilateral y simétrica (Guía de Práctica Clínica para las Distrofias Hereditarias de Retina, 2017). En estadios iniciales, los síntomas clásicos de la RP son la dificultad de adaptación a la oscuridad, ceguera nocturna y pérdida concéntrica del campo visual periférico, debido a la degeneración de los bastones. Los pacientes con RP normalmente presentan un fondo de ojo típico con palidez del nervio óptico, atenuación de los vasos retinianos y depósitos de pigmento en espículas en la retina periférica. Los depósitos de pigmento se forman cuando el pigmento del EPR migra a la retina neuronal por la muerte de los fotorreceptores. A medida que la enfermedad progresa, el campo visual se va reduciendo hasta alcanzar una típica visión en túnel debido a una progresiva constricción del campo visual (Hartong *et al.*, 2006). Al final de la enfermedad, la mayoría de los pacientes están abocados a la ceguera legal, definida por una agudeza visual $< 0,1$ y/o campo visual $< 10^\circ$.

Hoy en día, las actividades nocturnas se desarrollan en presencia de una gran intensidad de luz, lo que permite que los conos puedan activarse. Por ello, cuando un paciente con RP se da cuenta de que presenta ceguera nocturna, la alteración de los conos ya suele estar presente (Hartong *et al.*, 2006). Antes de manifestar disminución de la agudeza visual, los pacientes pueden perder hasta el 90% de los conos foveales (Geller & Sieving, 1993). En estos estadios finales de la enfermedad, los pacientes suelen manifestar dificultades en la lectura y la realización de las actividades diarias diurnas, lo que resulta más evidente cuando la agudeza visual cae por debajo de 0,5 (Grover *et al.*, 1999).

En la actualidad, el diagnóstico de RP se basa en la medición de parámetros oftalmológicos objetivos más que en los síntomas que el paciente presente. Las pruebas oftalmológicas a realizar en pacientes con sospecha de RP y/o posteriormente durante su seguimiento se detallan a continuación (Hartong *et al.*, 2006; Guía de Práctica Clínica para las Distrofias Hereditarias de Retina, 2017):

- **Adaptometría**, mide el tiempo de adaptación a la oscuridad, primer síntoma de alteración de los bastones.
- Medida del **campo visual** mediante el perímetro de Goldman o un analizador de campo Humphrey.

- **Agudeza visual** medida con optotipos. Puede permanecer normal incluso en pacientes con RP avanzada, debido a la preservación de campo visual central o puede perderse rápidamente con el curso de la enfermedad.
- **Electrorretinograma (ERG)**, mide la respuesta eléctrica de la retina en condiciones escotópicas y fotópicas, y con ello, proporciona información del estado funcional de los fotorreceptores. Los pacientes con RP presentan un ERG escotópico disminuido o ausente.
- **Oftalmoscopia**, consiste en la exploración del fondo de ojo (FO).
- **Test de visión de los colores**, mediante el test de Ishihara o Farnsworth 15D. Los pacientes con RP suelen tener una visión cromática normal, pero en estadios avanzados pueden presentar una deficiencia en la función de los conos azules.
- **Autofluorescencia**, para observar posibles cambios a nivel del EPR.
- **Tomografía de coherencia óptica (OCT)**, permite visualizar las distintas capas de la retina con una alta resolución.

Ante un paciente con síntomas de ceguera nocturna y sin hallazgos concluyentes en el FO, se recomienda realizar inicialmente una curva de adaptación a la oscuridad (Guía de Práctica Clínica para las Distrofias Hereditarias de Retina, 2017). En el caso de que se confirme la ceguera nocturna se debe realizar un ERG para valorar la respuesta escotópica de los bastones.

Dentro de la RP existen casos sindrómicos, que representan del 20-30% de los casos totales de RP, en los que los síntomas oculares aparecen acompañados de la afectación de otros órganos. Actualmente, se han descrito más de 30 síndromes diferentes. El síndrome más prevalente, representa el 15 al 20% de la RP, es el síndrome de Usher que se caracteriza por RP e hipoacusia neurosensorial, a veces con disfunción vestibular. El segundo síndrome más frecuente es el síndrome de Bardet-Biedl, que constituye el 5% de los casos de RP, se caracteriza por la asociación de RP con obesidad, polidactilia, alteraciones renales y retraso mental moderado (Ayuso & Millan, 2010).

1.4.2. Aspectos genéticos

La RP representa un modelo de extrema complejidad a nivel genético, para la que se han descrito casos asociados a cualquier patrón de herencia Mendeliana: autosómica recesiva (arRP), autosómica dominante (adRP) y ligada al cromosoma X (xLRP), así como a otras formas minoritarias como formas mitocondriales, digenismo, trialelismo e isodisomía uniparental

(Hartong *et al.*, 2006; Ayuso & Millan, 2010). Existen numerosos casos esporádicos de RP (sRP) y, aunque *a priori* no se puede inferir su patrón hereditario, suelen corresponder con formas recesivas (Wang *et al.*, 2005). En población española, los pacientes sRP representan el 40%, seguida de las formas recesivas (38%), dominantes (15%) y ligadas al cromosoma X (6%) (Ayuso *et al.*, 1995).

Actualmente, se han identificado más de 60 genes asociados a RP no sindrómica (RetNet). En general, la mayoría de estos genes explican un pequeño porcentaje de familias, con algunas excepciones, como el gen *RPGR*, responsable de >70 % de las formas ligadas al sexo (Hartong *et al.*, 2006), el gen *RHO* que representa el 20-25 % de los casos de adRP (Millá *et al.*, 2002), el gen *USH2A* que causa el 10% de las formas recesivas de RP (Blanco-Kelly *et al.*, 2015) o el gen *CRB1* que causa el 9% de las formas recesivas con inicio precoz (Corton *et al.*, 2013). Por ello, las mutaciones responsables de la RP son exclusivas de una única familia, ya que las mutaciones prevalentes o *hotspots* son infrecuentes (Ayuso & Millan, 2010). Esta elevada heterogeneidad genética y alélica dificulta el diagnóstico genético de la RP.

1.5. Retinosis pigmentaria autosómica dominante

El modelo de herencia dominante se caracteriza por una transmisión vertical, en donde ambos sexos tienen la misma probabilidad de ser afectados, los hermanos e hijos de un paciente diagnosticado de adRP tienen un riesgo del 50% de ser también afectados, por lo que las familias presentan miembros con afectación en cada generación. Habitualmente, se sospecha de adRP en aquellos casos familiares en los que existan tres o más generaciones de individuos afectados, con mujeres afectas y/o al menos dos generaciones si se observa transmisión varón-varón, lo que descarta formas ligadas al cromosoma X (Daiger *et al.*, 2014a). En población española, el 15% de los casos de RP no sindrómica corresponde a formas de adRP (Ayuso *et al.*, 1995).

Actualmente se han descrito 28 genes asociados a adRP (RetNet, revisado en Enero 2018) (Tabla 1). Las proteínas codificadas por estos genes están implicadas en muchos aspectos de la función del fotorreceptor, incluyendo la cascada de la fototransducción (*RHO*), su metabolismo (enzimas implicadas en rutas celulares importantes en el organismo como: *HK1* o *IMPDH1*; o la degradación proteica: *KLHL7*), genes implicados en el proceso de *splicing* que se expresan de forma ubicua (*PRPF3*, *PRPF8*, *PRPF31*, *SNRNP200*, *PRPF4* y *PRPF6*), factores de transcripción (*CRX*, *NRL*, *NR2E3*), proteínas estructurales localizadas en la membrana de los discos del fotorreceptor (*RHO*, *PRPH2*, *ROM1*) o en el cilio conector (*RP1*, *TOPORS*). En otros casos no se conoce la función de los genes asociados a adRP, tal como se describe en la Tabla 1.

Además de esta **elevada heterogeneidad** genética, las formas de adRP presentan una clínica variable caracterizada por una edad tardía de aparición de los síntomas, una progresión menos severa que en las formas recesivas y ligadas al cromosoma X, la posible existencia de **penetrancia incompleta** (se produce cuando un individuo es portador de una mutación pero no desarrolla la enfermedad) en algunas familias, especialmente asociada a algunos genes, así como una expresividad variable asociada a algunas mutaciones concretas (Kim *et al.*, 1995). Todo ello dificulta no sólo la identificación de variantes heterocigotas que expliquen la patología sino además la correcta interpretación de los estudios de segregación ante una variante nueva con predictores patogénicos. Así, puede ocurrir el caso de que en un individuo joven clínicamente sano portador de una variante probablemente patogénica asociado a adRP no haya manifestado la enfermedad aún, o incluso que ésta nunca vaya a ser desarrollada debido a que la mutación no presente penetrancia completa en dicho individuo.

Actualmente, las mutaciones en los genes conocidos permiten resolver entre 56-78% de los casos de adRP, dependiendo del tipo de prueba genética usada y la población estudiada (Daiger *et al.*, 2014a; Van Cauwenbergh *et al.*, 2017).

Tabla 1: Genes asociados a adRP. Tabla adaptada de Daiger *et al.*, 2014a.

Locus	Símbolo	Gen	Otras enfermedades	Función
1p31.2	<i>RPE65</i>	<i>Retinal pigment epithelium specific 65 kD protein</i>	arRP, arLCA	Ciclo visual
1q32.1	<i>ADIPOR1</i>	<i>Adiponectin receptor 1</i>	arRP, arRP sindrómica	Desconocida
1q21.2	<i>PRPF3</i>	<i>Pre-mRNA processing factor 3</i>	Ninguna	<i>Splicing</i>
1q22	<i>SEMA4A</i>	<i>Semaphorin 4A</i>	adDCB	Respuesta neuronal o inmune
2q11.2	<i>SNRNP200</i>	<i>Small nuclear ribonucleoprotein 200 kDa (U5)</i>	Ninguna	<i>Splicing</i>
2q37.1	<i>SPP2</i>	<i>Secreted phosphoprotein 2</i>	Ninguna	Renovación ósea
3q22.1	<i>RHO</i>	<i>Rhodopsin</i>	adCSNB, arRP	Fototransducción, estructura de la membrana del segmento externo
6p21.1	<i>GUCA1B</i>	<i>Guanylate cyclase activating protein 1B</i>	adMD	Fototransducción
6p21.1	<i>PRPH2</i>	<i>Peripherin 2</i>	adMD, RP digénica con <i>ROM1</i> ; adDM viteliforme del adulto, adCRD, distrofia coroidal areolar central (ad)	Estructura de la membrana del segmento externo
7p14.3	<i>RP9</i>	<i>RP9 protein or PIM1-kinase associated protein 1</i>	Ninguna	Proliferación de las células B
7p15.3	<i>KLHL7</i>	<i>Kelch-like 7 protein(Drosophila)</i>	Ninguna	Ruta de la ubiquitización
7q32.1	<i>IMPDH1</i>	<i>Inosine monophosphate dehydrogenase 1</i>	adLCA	Síntesis de las purinas o aminoácidos
8q12.1	<i>RP1</i>	<i>RP1 protein</i>	arRP	Estructura/función ciliar
9p21.1	<i>TOPORS</i>	<i>Topoisomerase I binding arginine/ serine rich protein</i>	Ninguna	Estructura/función ciliar
9q32	<i>PRPF4</i>	<i>Pre-mRNA processing factor 4</i>	Ninguna	<i>Splicing</i>
10q22.1	<i>HK1</i>	<i>Hexokinase 1</i>	Anemia hemolítica no esferocítica, neuropatía hereditaria recesiva	Metabolismo de la glucosa

Tabla 1: Genes asociados a adRP. Tabla adaptada de Daiger *et al.*, 2014a (Continuación).

Locus	Símbolo	Gen	Otras enfermedades	Función
10q24.32	<i>ARL3</i>	<i>ADP ribosylation factor like GTPase 3</i>	Ninguna	Estructura/función ciliar
11q12.3	<i>BEST1</i>	<i>Bestrophin 1</i>	ar bestrofinopatía, arRP, adMD tipo Best	Ciclo visual
11q12.3	<i>ROM1</i>	<i>Retinal outer segment membrane protein 1</i>	RP digénica con <i>PRPH2</i>	Estructura de la membrana del segmento externo
14q11.2	<i>NRL</i>	<i>Neural retina lucine zipper</i>	arRP	Factor de transcripción
14q24.1	<i>RDH12</i>	<i>Retinol dehydrogenase 12</i>	arLCA	Ciclo visual
15q23	<i>NR2E3</i>	<i>Nuclear receptor subfamily 2 group E3</i>	arRP, síndrome de Goldmann–Favre (ar), síndrome de conos S aumentados (ar)	Factor de transcripción
17p13.3	<i>PRPF8</i>	<i>Pre-mRNA processing factor 8</i>	Ninguna	<i>Splicing</i>
17q23.2	<i>CA4</i>	<i>Carbonic anhydrase IV</i>	Ninguna	Estructura/función de los capilares de la retina
17q25.3	<i>FSCN2</i>	<i>Retinal fascin homolog 2, actin bundling protein</i>	adMD	Estructura/función ciliar
19q13.32	<i>CRX</i>	<i>Cone–rod otx-like photoreceptor homeobox transcription factor</i>	arLCA, adLCA, adDCB	Factor de transcripción
19q13.42	<i>PRPF31</i>	<i>Pre-mRNA processing factor 31</i>	Ninguna	<i>Splicing</i>
20q13.33	<i>PRPF6</i>	<i>Pre-mRNA processing factor 6</i>	Ninguna	<i>Splicing</i>

Leyenda: ad: autosómico dominante; ar: autosómico recesivo; CSNB: ceguera nocturna congénita estacionaria; DCB: Distrofia de Conos-Bastones; DM: Distrofia Macular; LCA: *Leber Congenital Amaurosis*; RP: Retinosis Pigmentaria.

1.6. El gen *PRPF31* y la penetrancia incompleta

El gen *PRPF31* (*pre-mRNA processing factor 31*; NM_015629), está localizado en el locus 19q13.42 y está compuesto por 14 exones. Representa uno de los genes más prevalentes en la adRP, con una frecuencia de aproximadamente el 8% (Al-Magthteh *et al.*, 1996; Vithana *et al.*, 2001; Ferrari *et al.*, 2011). El gen *PRPF31* es de expresión ubicua, incluyendo tejidos neuronales, cerebro y retina (Vithana *et al.*, 2001). Codifica para una proteína de *splicing* de 499 aminoácidos, muy conservada evolutivamente desde levaduras a mamíferos (Makarova *et al.*, 2002), con un dominio Nop de unión a ribonucleoproteínas. PRPF31 presenta un papel fundamental en la formación de tri-ribonucleoproteína nuclear U4/U6*U5, molécula clave del espliceosoma (Liu *et al.*, 2007). El espliceosoma es un complejo ribonucleoproteico con actividad nucleasa y ligasa que está formado por moléculas snRNAs (*small nuclear RNAs*) y proteínas, que es responsable del corte y empalme del mRNA durante el proceso de *splicing* (Tarn & Steitz, 1997) (Figura 7). Dado que PRPF31 participa en la eliminación de los intrones de los pre-mRNAs durante el proceso de *splicing*, el mecanismo patogénico asociado a adRP mediado por mutaciones en este gen está asociado a la desregulación en la formación del espliceosoma (Makarova *et al.*, 2002). A pesar de su expresión ubicua, las mutaciones en este gen solo se han asociado a adRP, probablemente debido a que la retina presenta la mayor tasa de expresión génica de todos los tejidos del cuerpo. Por ello, los fotorreceptores requieren una gran demanda de componentes funcionales de *splicing* (Cao *et al.*, 2011; Tanackovic *et al.*, 2011), por lo que la haploinsuficiencia de este factor de *splicing* conduce a la muerte del fotorreceptor (Cao *et al.*, 2011).

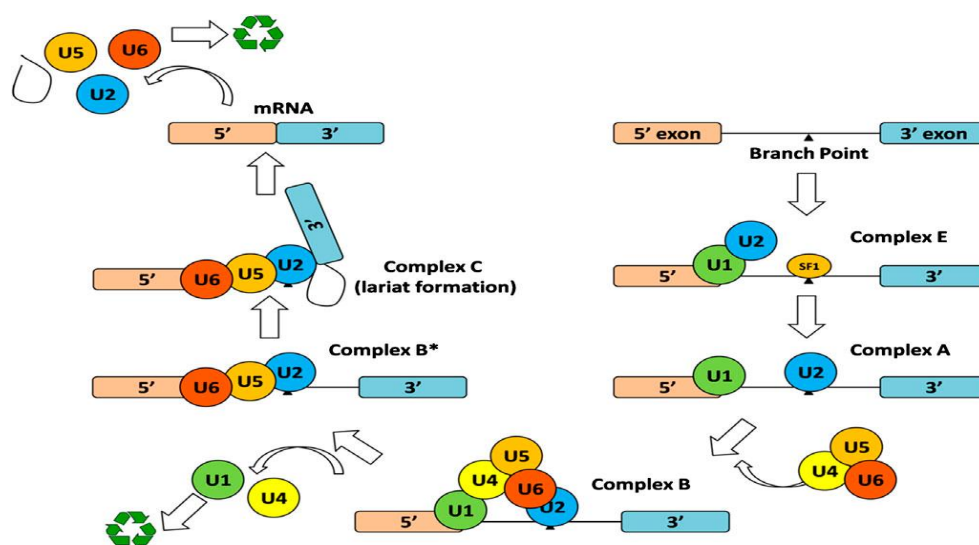


Figura 7. Representación esquemática del proceso de *splicing* (Rose *et al.*, 2016).

La **penetrancia incompleta** representa un mecanismo frecuente en familias con mutaciones en *PRPF31* (Daiger *et al.*, 2014a). La presencia de portadores que se mantienen asintomáticos incluso a edades avanzadas, parece venir explicada por un aumento de la expresión del alelo *wild-type* que permite compensar la disminución de la expresión de *PRPF31* asociada al alelo mutado (McGee *et al.*, 1997). Se ha descrito que *PRPF31* tiene una expresión génica que sigue una distribución continua en la población general, con un umbral límite que asegura la función normal del mecanismo de *splicing* y, a su vez de la retina (Figura 8). En este sentido, se ha descrito que los portadores asintomáticos normalmente presentan niveles altos de expresión de *PRPF31* por encima del umbral, y por tanto no desarrollarían la enfermedad (Rose & Bhattacharya, 2016).

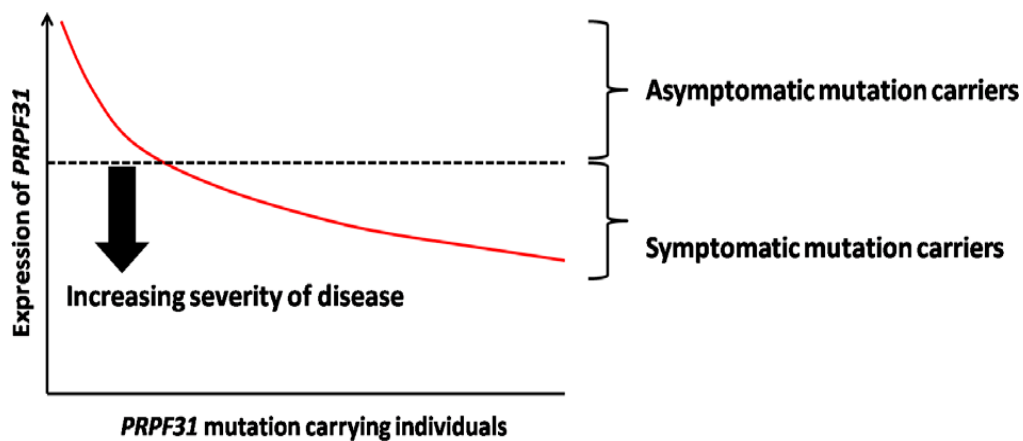


Figura 8. Distribución continua de la expresión de *PRPF31*. El nivel de expresión del gen *PRPF31* en población general muestra una distribución continua, representada con una línea roja. Existe un umbral para la función normal de la retina (línea discontinua). Cuando un portador de mutación en *PRPF31* tiene un nivel residual de proteína por debajo de este umbral, presentará síntomas (Rose *et al.*, 2016).

Se ha postulado que este fenómeno está relacionado con la regulación transcripcional de *PRPF31* mediada por diferentes factores transactivadores (Rose *et al.*, 2014) (Figura 9). Entre los posibles genes candidatos que han sido identificados, varios trabajos apuntan a que el gen *CNOT3*, cuya expresión es inversamente proporcional a la de *PRPF31*, podría actuar como un gen modificador de la penetrancia incompleta en portadores de mutaciones en *PRPF31* (Venturini *et al.*, 2012; Rose *et al.*, 2014) (Figura 9B). La supresión *in vitro* de *CNOT3*, mediada por su inhibición con siRNAs específicos, conlleva al aumento de la expresión de *PRPF31*, lo que indicaría que *CNOT3* es un regulador negativo de *PRPF31* (Venturini *et al.*, 2012). *CNOT3* modularía la transcripción génica a través de su unión directa al promotor de *PRPF31* (Hu *et al.*, 2009; Rose *et al.*, 2016).

Un segundo mecanismo modulador, propuesto recientemente, vendría mediado por un minisatélite MSR1 (*minisatellite repeat element*) localizado *upstream* al promotor de *PRPF31* (Rose *et al.*, 2016). Los elementos repetidos MSR1 son minisatélites de 36-38pb presentes a lo largo de todo el genoma, pero que se encuentran en una gran proporción en el cromosoma 19. Los elementos MSR1 tienen una conservación específica de *locus*, por ello se ha postulado que deben de tener un papel funcional, muy probablemente modulando la expresión génica (Bhavsar *et al.*, 1996; Rose *et al.*, 2016). Este MSR1 presenta dos formas alélicas, con 3 o 4 copias, que conducen a un efecto inverso sobre la activación transcripcional de *PRPF31* (Rose *et al.*, 2016) (Figura 9B). Mientras que el alelo de 3 copias conduce a una disminución en la expresión de *PRPF31*, la forma de 4 copias, con una MAF (*minor allele frequency*) de 15,4% en caucásicos, aumenta la actividad transcripcional del gen (Rose *et al.*, 2016). Estos mismos autores describen que ninguno de los pacientes afectados de RP, de una cohorte de 42 individuos sintomáticos y 29 asintomáticos, portadores de mutaciones en *PRPF31*, presentaban el alelo de 4 copias. Además, en cohortes de origen chino y japonés, donde la penetrancia incompleta para *PRPF31* no es frecuente, la frecuencia del alelo de 4 copias es muy inferior a la descrita en población caucásica. Todas estas evidencias indicarían que el alelo de 4 copias para este MSR1 podría actuar como un factor protector frente a mutaciones en *PRPF31* situadas en *cis* con el alelo *wild-type*, lo que determinaría que estos portadores se mantuvieran asintomáticos (Rose *et al.*, 2016).

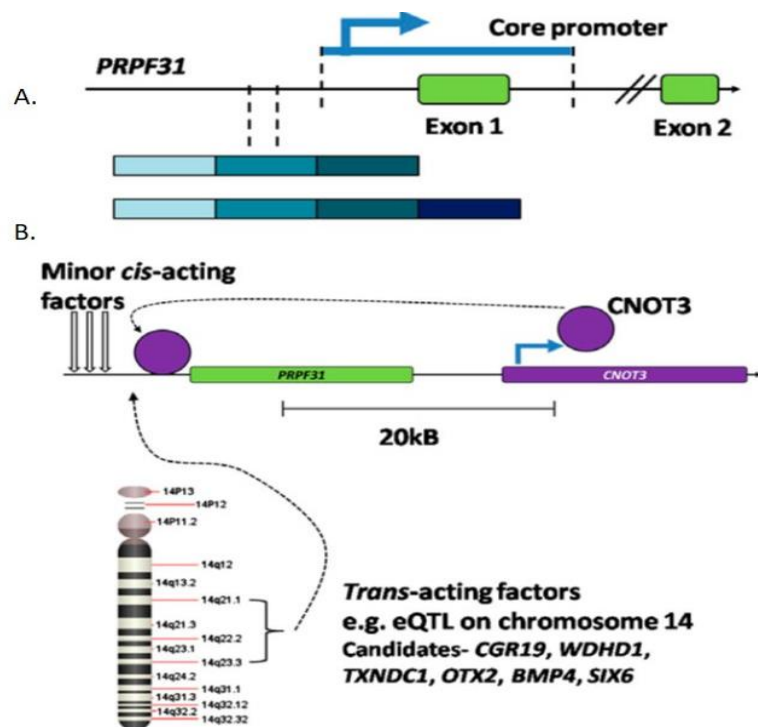


Figura 9. Control de la expresión génica de *PRPF31*, mediada por: A. el número de copias del minisatélite MSR1; B. el gen *CNOT3* y otros factores transactivadores (Rose *et al.*, 2016).

1.7. Estudio molecular de la adRP

La elevada heterogeneidad clínica, genética y alélica de las formas dominantes de RP, hace que el diagnóstico molecular de la enfermedad sea un proceso costoso a nivel económico y de recursos. El estudio genético de la adRP ha ido cambiando a medida que aumentaba el conocimiento sobre los genes identificados, los mecanismos moleculares subyacentes y evolucionaban de forma paralela las técnicas usadas en el cribado molecular de las mutaciones.

Inicialmente, la combinación de **técnicas clásicas de cribado indirecto** como *SSCP* (*Single Stranded Conformational Polymorphism*) o *DGGE* (*Denaturing Gradient Gel Electrophoresis*) junto con la secuenciación Sanger, se aplicaban para detectar los cambios puntuales más prevalentes asociados a adRP (*RHO*, *PRPH2*, ...) (Martínez-Gimeno *et al.*, 2003). De forma alternativa, en la última década se diseñaron arrays de genotipado, basados en la tecnología *APEX* (*Arrayed Primer Extension*), que permiten detectar mutaciones conocidas en genes específicos (Blanco-Kelly *et al.*, 2012). En su última versión, el chip de genotipado específico para adRP detecta >400 mutaciones en 16 genes asociados a adRP. La aplicación de esta técnica permitió a nuestro grupo una tasa diagnóstica del 15% en una cohorte de 139 familias (Blanco-Kelly *et al.*, 2012).

Estas técnicas hacían que el diagnóstico molecular de las DR fuese muy complejo, laborioso, con un elevado gasto de dinero y tiempo. La aparición de nuevas **tecnologías de NGS**, que permiten el análisis directo de la secuencia genómica de una forma rápida y coste efectiva, ha supuesto una auténtica revolución en el estudio molecular de enfermedades genéticas. A lo largo de los últimos años se han desarrollado distintas plataformas y estrategias de análisis que han ido reduciendo el coste, complejidad de los protocolos y el análisis bioinformático de los datos. Esto ha permitido una rápida implementación de la NGS en la rutina diagnóstica de la RP (Sorrentino *et al.*, 2016), y así como su uso para la identificación de nuevos genes (Avila-Fernandez *et al.*, 2016; Corton *et al.*, 2016). Por ello, esta tecnología se está imponiendo como primer abordaje para todas aquellas patologías con elevado número de genes asociados, como ocurre en la RP (Sorrentino *et al.*, 2016). Existen varias aproximaciones que pueden aplicarse en el estudio de la adRP, en función del tamaño de la región de interés a ser secuenciada, mediante una aproximación de amplicones de PCR o de captura utilizando sondas de hibridación (Mamanova *et al.*, 2010; Kozarewa *et al.*, 2015). A continuación, se describen los distintos tipos de NGS:

- **Panel de genes:** permite la secuenciación simultánea de un grupo de genes asociados a la patología. Su aplicación a la adRP ha permitido aumentar la tasa diagnóstica en trabajos previos a esta Tesis Doctoral (Fernandez-San Jose *et al.*, 2015a).
- **Exoma clínico** (www.illumina.com/products/by-type/clinical-research-products/trusight-one.html): en el que se capturan y pueden secuenciarse todos los exones codificantes de aquellos genes previamente asociados a patología en OMIM (*Online Mendelian Inheritance in Man*, www.omim.org).
- **Exoma completo** (*Whole Exome Sequencing*, WES): se secuencian todas las regiones codificantes (o exoma) del genoma (Bao *et al.*, 2014; Almoguera *et al.*, 2015).
- **Genoma completo** (*Whole Genome Sequencing*, WGS): se produce la secuenciación de todo el genoma, incluidas las regiones intrónicas y reguladoras (Nishiguchi *et al.*, 2013).

Además del estudio de variantes puntuales, en los últimos años se ha puesto de manifiesto que las variación en el número de copia (CNV), tanto deleciones como duplicaciones, representan una importante fuente de causalidad en las distrofias de retina (Bujakowska *et al.*, 2017; Van Schil *et al.*, 2017). Para su estudio se pueden utilizar distintas estrategias, las más tradicionales basadas en hibridación de sondas como el MLPA (*Multiplex Ligation Probe Amplification*) o los arrays de CGH (*Comparative Genomic Hybridization array*). Recientemente, el desarrollo de nuevos algoritmos bioinformáticos permite el análisis de CNVs a partir de los datos generados mediante NGS (Ellingford *et al.*, 2017).

Debido a que la mayoría de los estudios realizados en la actualidad en pacientes con DR se suelen realizar mediante secuenciación masiva utilizando paneles de genes, exoma o WGS, no se establecen cohortes “*a priori*” según fenotipo y patrón de herencia para el estudio de los diferentes subtipos (Glockle *et al.*, 2014; Oishi *et al.*, 2014; Bravo-Gil *et al.*, 2016). Por ello, el rendimiento diagnóstico específico de la adRP, así como la composición génica, no está bien establecido, siendo estos aspectos cruciales para un correcto asesoramiento genético de estas familias.

La aplicación de forma sistemática de las nuevas técnicas de NGS a una cohorte de familias con adRP, previamente estudiadas con métodos clásicos, o bien recogidas de forma prospectiva, permitirán establecer un aumento de la tasa de caracterización molecular, lo que permite un adecuado diagnóstico y consejo genético, además de la futura inclusión de los pacientes en terapias y ensayos clínicos.

Objetivos

Objetivo general

El objetivo principal de esta Tesis doctoral ha sido profundizar en el conocimiento de las bases moleculares de la retinosis pigmentaria autosómica dominante (adRP) en España, mediante su caracterización genética, fenotípica y epidemiológica, con el fin de elaborar un nuevo algoritmo que nos permita agilizar, optimizar y disminuir el coste de su diagnóstico genético.

Objetivos específicos

1. Caracterización clínica y genética de familias españolas con adRP, utilizando nuevas técnicas de secuenciación masiva para la búsqueda de nuevas mutaciones, con la finalidad de mejorar su diagnóstico y consejo genético.
2. Estudio sobre la implicación de las CNVs en la adRP.
3. Comparación de la eficacia del diagnóstico molecular de adRP, utilizando herramientas convencionales, respecto del uso de las nuevas técnicas de secuenciación masiva.
4. Análisis epidemiológico molecular en pacientes españoles con adRP: análisis de la tasa de caracterización molecular actual e identificación de los genes y mutaciones más prevalentes.
5. Estudio del papel del gen *PRPF31* en familias españolas con sospecha de adRP.
6. Análisis de correlación genotipo-fenotipo en pacientes con mutaciones en el gen *PRPF31*.

Material y Métodos

3.1. Sujetos y muestras

Los pacientes objeto de estudio de esta Tesis Doctoral se reclutaron en el Servicio de Genética del Hospital Universitario Fundación Jiménez Díaz (FJD) a lo largo de 26 años (1991-2017), tras haber sido remitidos para la realización de estudios genéticos de retinosis pigmentaria.

El **diagnóstico clínico** de RP se realizó en base a exámenes oftalmológicos y/o neurofisiológicos, incluyendo estudios de agudeza visual, campimetría, fundoscopia y ERG, siguiendo los criterios clínicos establecidos (Guía de Práctica Clínica para las Distrofias Hereditarias de Retina, 2017).

La **clasificación genética** de un patrón de herencia autosómica dominante se realizó por la observación de la patología, en al menos dos generaciones consecutivas de individuos afectos.

En este trabajo, se seleccionaron pacientes y/o familias que cumplieron los siguientes criterios:

i) sospecha inicial de RP tras revisión de su anamnesis e informes oftalmológicos y ii) presentación de una historia familiar de RP, compatible con una herencia autosómica dominante. Se excluyeron pacientes que presentaban síntomas y signos clínicos de RP síndromica y los pacientes con árbol genealógico compatible con herencia autosómica recesiva o ligada al cromosoma X.

Un total de **258 familias españolas con diagnóstico clínico inicial de adRP** fueron incluidas en este estudio, tanto los casos índices, que fueron analizados con los estudios genéticos detallados a continuación, así como otros familiares afectos y sanos, en los que se realizaron estudios de segregación y/o estudios de exoma completo. Adicionalmente, se utilizó una cohorte de 132 individuos españoles sanos, sin síntomas ni signos de enfermedades oculares, como controles en varios de los estudios realizados. Esta cohorte fue reclutada a través del Banco de Sangre de la FJD y las muestras fueron anonimizadas, una vez recogidas. Todos los individuos fueron incluidos en este estudio tras la firma de un consentimiento informado. Este estudio fue revisado y aprobado por el Comité de Ética de nuestro hospital y se realizó de acuerdo con los principios de la Declaración de Helsinki.

Se utilizaron muestras de ADN y ARN obtenidas de forma prospectiva específicamente para este estudio o procedentes de la colección de muestras del Biobanco del Hospital Universitario FJD.

Las **muestras de ADN genómico** se extrajeron a partir de sangre periférica usando dos tipos de extractores de ADN automatizados: *BioRobot EZ1* (QIAGEN, Hilden, Germany) o *MagNA Pure Compact System* (Roche Applied Science, Penzberg, Germany).

Las **muestras de ARN** fueron extraídas a partir de sangre total humana usando el kit *PAXgene Blood RNA* y el extractor automatizado *QIAcube* (QIAGEN, Hilden, Germany), siguiendo las recomendaciones y los protocolos de la casa comercial.

3.2. Cribado mutacional de variantes de un solo nucleótido (SNV)

El cribado mutacional del caso índice de cada familia se realizó usando diferentes técnicas moleculares a lo largo del tiempo, como **primer enfoque diagnóstico**:

i) **Abordaje clásico** basado en el cribado directo y/o indirecto de los genes más prevalentes asociados a adRP, principalmente *RHO* (Fernandez-San Jose *et al.*, 2015b), usando SSCP, DGGE (Martinez-Gimeno *et al.*, 2003) y/o secuenciación Sanger, en combinación con el genotipado de mutaciones conocidas en 16 genes mediante un chip comercial de adRP (Asper Biotech, Tartu, Estonia) (Blanco-Kelly *et al.*, 2012).

ii) **Secuenciación masiva** mediante la aproximación de exoma clínico, para estudiar 32 nuevas familias recopiladas durante el desarrollo de esta Tesis Doctoral, y que no fueron analizadas previamente mediante el abordaje clásico. Para el análisis de estos pacientes se utilizó un **panel virtual** de 74 genes asociados a DR (Tabla 2).

Adicionalmente, un total de 105 familias no caracterizadas con el abordaje clásico se analizaron mediante distintas aproximaciones de NGS, tanto con una estrategia basada en la captura de paneles personalizados de genes asociados a distintas DR (Tabla 2), como con secuenciación del exoma completo (Almoguera *et al.*, 2015).

El **listado de genes** capturados mediante las distintas estrategias de NGS utilizadas en este trabajo se detalla en la Tabla 2. En general, se pudieron analizar 24 de los 28 genes cribados asociados a adRP (RetNet). En el trabajo no se abordó el estudio de los genes *ADIPOR1*, *ARL3*, *PRPF4* y *SPP2*, que han sido recientemente asociados a adRP y no estaban incluidos en los paneles personalizados ni son capturados mediante la aproximación utilizada de exoma clínico. El gen *HK1* solo puso ser analizado mediante exoma clínico en 32 pacientes. Además, estos paneles permitieron analizar otros genes de DR no sindrómica, principalmente asociados a formas de arRP y xLRP.

A continuación, se describen las principales aproximaciones de NGS utilizadas en esta Tesis Doctoral:

Tabla 2. Listado de genes asociados a distrofias de retina y estudiados en el presente estudio. Se muestran los distintos genes incluidos en cada tipo de tecnología de NGS: HaloPlex (75 genes), MIPs (108 genes) o exoma clínico (74 genes) (TruSight One –Illumina-). En negrita 24 genes cribados de los 28 genes asociados a adRP (RetNet).

Gen	HaloPlex	MIPs	Exoma	Gen	HaloPlex	MIPs	Exoma
ABCA4	x	x	x	MPDZ	x		
ABHD12	x	x	x	NEK2		x	
ADAM9		x		NMNAT1	x	x	x
ADAMTS18	x	x	x	NR2E3	x	x	x
AIPL1	x	x	x	NRL	x	x	x
ARL2BP		x		OFD1	x	x	x
ARL6		x		OTX2		x	
BBS1	x	x	x	PDE6A	x	x	x
BBS2		x		PDE6B	x	x	x
BEST1	x	x	x	PDE6C		x	
C1QTNF5		x	x	PDE6G	x	x	x
C21orf2		x		PDE6H		x	
C2orf71	x	x	x	PDZD7		x	
C8orf37	x	x	x	PITPNM3		x	
CA4	x	x	x	PRCD	x	x	x
CABP4	x	x	x	PROM1	x	x	x
CACNA1F		x		PRPF3	x	x	x
CACNA2D4		x		PRPF31	x	x	x
CDHR1		x		PRPF6	x	x	x
CEP290	x	x	x	PRPF8	x	x	x
CERKL	x	x	x	PRPH2	x	x	x
CHM	x	x	x	RAB28		x	
CLN3		x		RAX2		x	
CLRN1	x	x	x	RBP3	x	x	x
CNGA1	x	x	x	RD3	x	x	x
CNGA3		x		RDH12	x	x	x
CNGB1	x	x	x	RDH5		x	x
CNGB3		x		RGR	x	x	x
CNNM4		x		RHO	x	x	x
CRB1	x	x	x	RIMS1		x	
CRX	x	x	x	RLBP1	x	x	x
CYP4V2	x	x	x	ROM1	x	x	x
DHDDS	x	x	x	RP1	x	x	x
DTHD1		x		RP1L1		x	
EMC1		x		RP2	x	x	x
EYS	x	x	x	RP9	x	x	x
FAM161A	x	x	x	RPE65	x	x	x
FSCN2	x	x	x	RPGR	x	x	x
GPR125		x		RPGRIP1	x	x	x
GUCA1A		x		SAG	x	x	x
GUCA1B	x	x	x	SAMD11	x		
GUCY2D	x	x	x	SEMA4A	x	x	x
HK1			x	SNRNP200	x	x	x
IDH3B	x	x	x	SPATA7	x	x	x
IMPDH1	x	x	x	TOPORS	x	x	x
IMPG1	x	x		TTC8	x	x	x
IMPG2	x	x	x	TULP1	x	x	x
IQCB1	x	x	x	UNC119		x	
ITM2B		x		USH1C		x	
KCNJ13	x	x	x	USH2A	x	x	x
KCNV2		x		VCAN		x	
KIAA1549		x		WDR19		x	
KLHL7	x	x	x	ZNF408	x	x	
LCA5	x	x	x	ZNF513	x	x	x
LRAT	x	x	x	ZNF408	x	x	
MAK	x	x	x				
MERTK	x	x	x				

3.2.1. Panel customizado de genes de RP mediante HaloPlex

Se trata de un panel customizado de 75 genes conocidos que causan DR, principalmente RP y LCA (Tabla 2), que fue desarrollado de forma personalizada usando la tecnología de captura HaloPlex (Agilent Technologies, Inc., Santa Clara, CA, USA), tal como se describe en las Tesis Doctorales de Patricia Fernández San José y Raquel Pérez Carro (Fernández-San-José, 2017; Pérez-Carro, 2017). Para ello, se diseñaron 13.676 amplicones, para 1.127 regiones diana, incluyendo exones codificantes y no codificantes, y 20 pb en las regiones intrónicas de *splicing*, usando la herramienta de *SureDesign* (Agilent Technologies, Inc.). El diseño final, presentaba un tamaño de 352 kb y cubría el 99,1% de las regiones diana.

En total, se analizaron 68 casos índice que no habían previamente sido caracterizados siguiendo el algoritmo "clásico" de diagnóstico genético como primer abordaje. Durante este trabajo de Tesis Doctoral, se reanalizaron 59 muestras cuyo análisis fue descrito en un artículo publicado en 2015 (Fernandez-San Jose *et al.*, 2015a), así como otros 9 nuevos casos con adRP, procesados posteriormente a dicho trabajo.

La captura de ADN diana se realizó de acuerdo con el sistema de enriquecimiento HaloPlex adaptado a la tecnología de secuenciación por síntesis (SBS, *sequencing by synthesis*) de Illumina, siguiendo el protocolo de la casa comercial (protocolo versión D3, December 2012; Agilent Technologies) (Figura 10). En este sistema, 450 ng de ADN fueron digeridos con una combinación de 8 enzimas de restricción. Los fragmentos digeridos fueron hibridados con sondas biotiniladas, diseñadas de forma específica para que hibriden a ambos extremos de las secuencias diana, con el fin de producir su circularización. Estas moléculas circulares contienen la región de interés, los adaptadores y los códigos de barra para permitir un análisis multiplexado de 11 muestras por carrera. Las regiones capturadas se purificaron mediante bolas magnéticas unidas a estreptavidina (*Dynabeads*, Invitrogen), fueron amplificadas mediante PCR y purificadas dos veces mediante AMPure XP beads (Beckman Coulter, Brea, CA, USA). Las librerías fueron secuenciadas en 6 carreras independientes utilizando un secuenciador MiSeq de Illumina (Illumina, San Diego, California, USA) generando lecturas *paired-end* de 2x150pb, con una profundidad media de >340X.

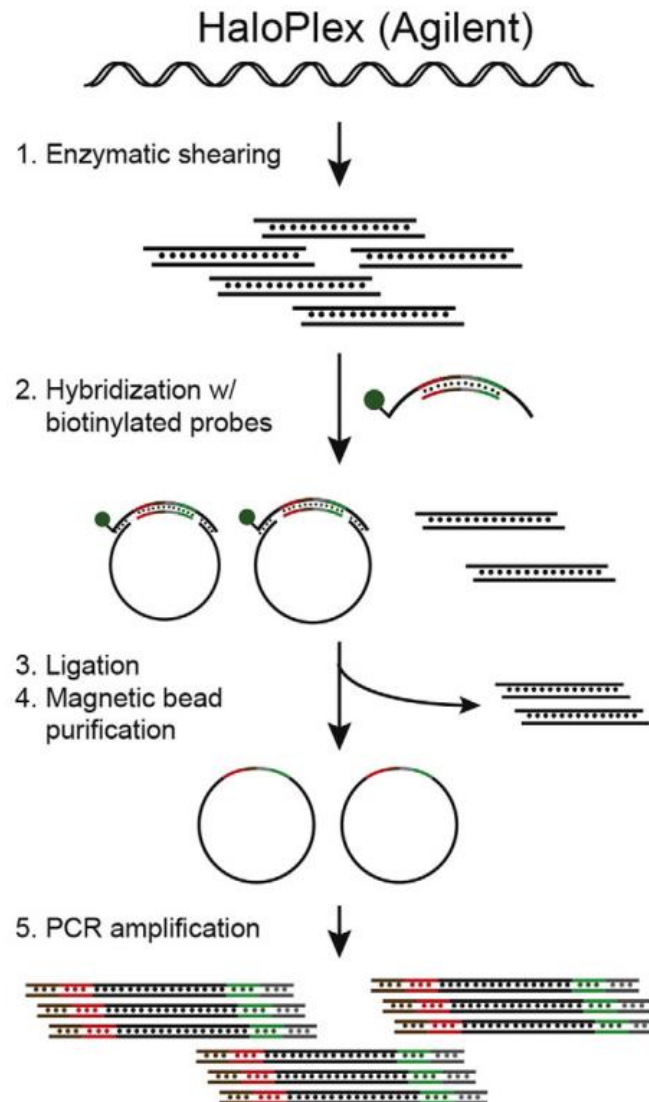


Figura 10. Representación esquemática de la preparación de las librerías del método de captura basado en HaloPlex. Imagen adaptada de Khodakov *et al.*, 2016.

El posterior análisis bioinformático de los datos de secuenciación se realizó con distintas estrategias. Primero, se utilizó la plataforma comercial *DNAexus* (www.DNAexus.com), utilizando una *pipeline* específicamente desarrollada para el análisis de librerías de HaloPlex (Fernandez-San Jose *et al.*, 2015a). En esta *pipeline* se utilizaron los algoritmos *Burrows-Wheeler Aligner* (BWA)-MEM versión 0.7.5a (Li & Durbin, 2010) durante el mapeo de las secuencias al genoma de referencia (hg19), *Genome Analysis Toolkit* (GATK) - lite version 2.3-9 (McKenna *et al.*, 2010) para la identificación (*calling*) de variantes tipo SNVs, y finalmente las herramientas *GATK Variant Annotator* y *ENSEMBL Variant Effect Predictor* v72 (McLaren *et al.*, 2010) fueron utilizadas en la anotación de las variantes en un archivo *vcf*, que fue utilizado posteriormente para la priorización de variantes potencialmente patogénicas.

En un segundo reanálisis bioinformático de las muestras realizado durante esta Tesis Doctoral, se utilizó un nuevo *pipeline* implementado en nuestro laboratorio (Figura 11), donde se utilizan versiones más actualizadas de las herramientas para el alineamiento y *calling*, que las previamente implementadas en el *pipeline* de DNANexus, así como nuevos programas para la anotación de las variantes (*Annovar*), nuevos predictores de patogenicidad (como M-CAP, CADD o Spidex) y nuevas bases de datos de frecuencias alélicas (*gnomAD*). Una vez analizadas, se reevaluaron las variantes previamente consideradas como patogénicas y la posible existencia de falsos negativos en el análisis previo, siguiendo los criterios descritos en el Apartado 3.2.5.

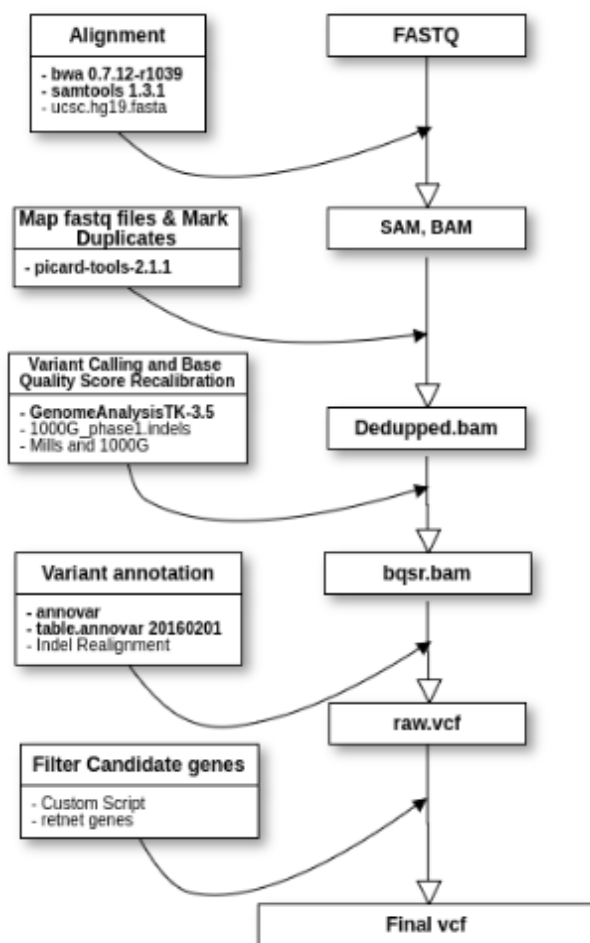


Figura 11. Pipeline implementado en nuestro laboratorio. Con esta pipeline se genera un archivo *vcf* a partir de los datos crudos del secuenciador.

3.2.2. Panel de genes de DR mediante la tecnología MIPs

De forma similar a HaloPlex, la tecnología MIPs (*Molecular inversion probes*), se trata de una técnica mixta a la hora de realizar el enriquecimiento, que se realiza mediante captura seguida de una amplificación por PCR. A diferencia de la tecnología anterior, esta tecnología no está disponible de forma comercial y para su implementación se utilizó la aproximación de O’Roak y

colaboradores (O'Roak *et al.*, 2012), descrita en la Figura 12. Para ello, se diseñaron 6.200 sondas específicas a lo largo de más de 1.500 exones de 108 genes (Tabla 2), previamente asociados a DR no sintomática (RetNet, diseño 2015), que cubría el 95% de las regiones diana.

Tanto el diseño, optimización de las sondas, preparación de las librerías, secuenciación y análisis bioinformático se llevaron a cabo en el laboratorio del Dr. Frans Cremers del *Radboud University Medical Center* (Nijmegen, Holanda), dentro del marco de un proyecto europeo colaborativo (ERDC4000) del Consorcio Europeo de Enfermedades de la Retina (ERDC), en donde se analizaron más de 4000 muestras de pacientes con distintas DR procedentes de 8 países europeos. Las librerías se secuenciaron en grupos de 95 muestras en un secuenciador NextSeq500 (Illumina) para obtener secuencias *paired-end* de 150 bp con cobertura media de 500X.

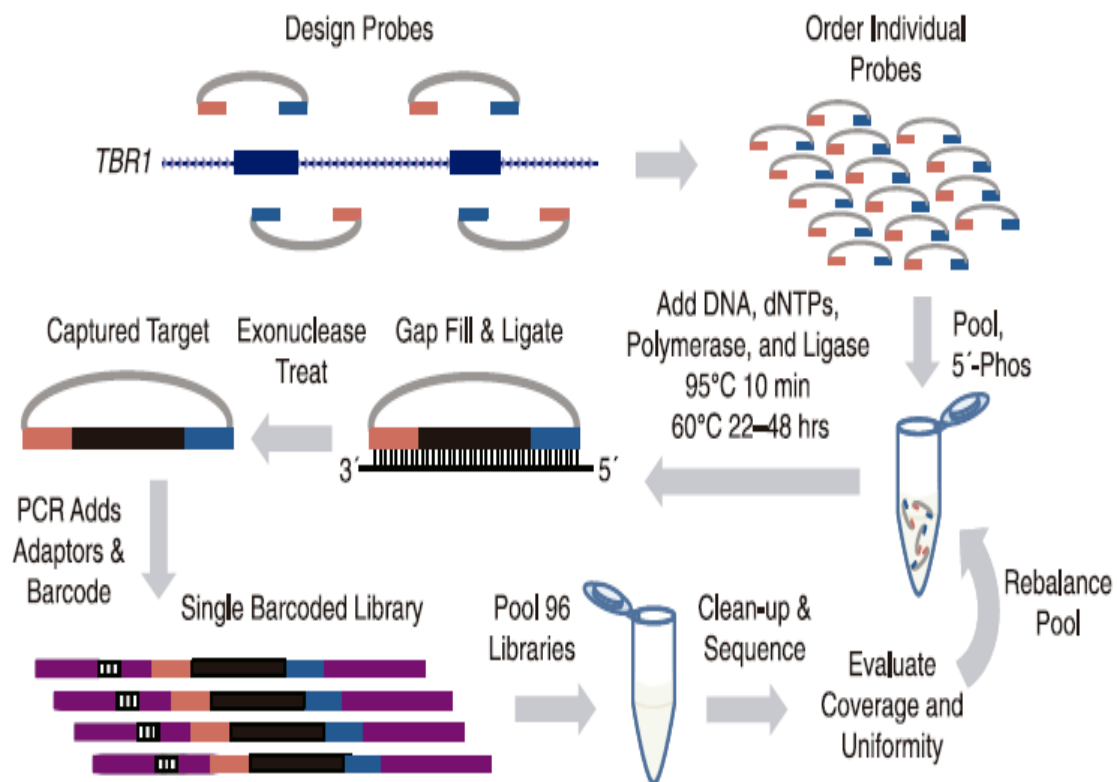


Figura 12: Representación esquemática del diseño y flujo de trabajo del método de captura basado en MIPs. Imagen adaptada de O'Roak *et al.*, 2012.

El análisis bioinformático fue realizado utilizando un *pipeline* basado en procedimientos estándares que fue específicamente adaptado para el análisis de esta estrategia de MIPs por el Servicio de Bioinformática del *Radboud University Medical Center* (Nijmegen, Holanda). La priorización e interpretación de la patogenicidad de las variantes en las familias en estudio fue realizada en los laboratorios de origen de las muestras.

En esta Tesis Doctoral se incluye el análisis mediante esta tecnología de 31 familias con adRP que no fueron caracterizadas previamente tras genotipado con el chip de adRP y/o secuenciación del gen *RHO* principalmente.

3.2.3. Secuenciación del exoma completo

Seis de las familias incluidas en este trabajo fueron previamente analizadas mediante WES en colaboración con el grupo del Dr. Hakon Hakonarson del *Center for Applied Genomics* del *Children's Hospital of Philadelphia* (Almoguera *et al.*, 2015). Las librerías de 38 individuos, tanto afectados como sanos, fueron preparadas usando el kit *Agilent Sure Select Human All Exon* versión 2 (Agilent Technologies, Santa Clara, CA, USA) y se analizaron en un secuenciador HiSeq 2000 (Illumina, San Diego, CA, USA), siguiendo los protocolos de las casas comerciales.

El análisis bioinformático fue realizado utilizando BWA (Li & Durbin, 2009) para el alineamiento de las secuencias al genoma de referencia (UCSC hg19), GATK versión 1.4 (McKenna *et al.*, 2010) para el *calling* de las variantes y la herramienta ANNOVAR (Wang *et al.*, 2010) para la anotación de las mismas (Almoguera *et al.*, 2015).

3.2.4. Exoma clínico

Treinta y dos familias fueron analizadas utilizando la aproximación de exoma clínico mediante la tecnología de *TruSight One* (Illumina), que permite capturar 12Mb de secuencia genómica de ~62.000 exones de 4.813 genes asociados a fenotipo clínico en OMIM y HGMD. Esta aproximación utiliza la tecnología Nextera para la preparación de librerías que combina una fragmentación del ADN de partida mediada por transposones ("tagmentación") con un posterior enriquecimiento de secuencias mediada por sondas de hibridación de 80 mer (Figura 13). Las librerías se prepararon de acuerdo a las instrucciones del fabricante y se secuenciaron en grupos de 24 muestras en la plataforma NextSeq500 (Illumina) utilizando el kit *NextSeq HighOutput*. El análisis bioinformático se realizó utilizando la plataforma online *BaseSpace* con la app *BWA Enrichment* (versión 2.1.2) de esta casa comercial, que para el alineamiento de las secuencias al genoma de referencia (hg19) y el *calling* de variantes tipo SNVs utiliza los algoritmos BWA y GATK, respectivamente. Posteriormente, la anotación y priorización de variantes fue realizado con el programa comercial *VariantStudio* (Illumina). Durante este análisis se priorizaron variantes potencialmente patogénicas en un subgrupo de 74 genes asociados a DR (Tabla 2).

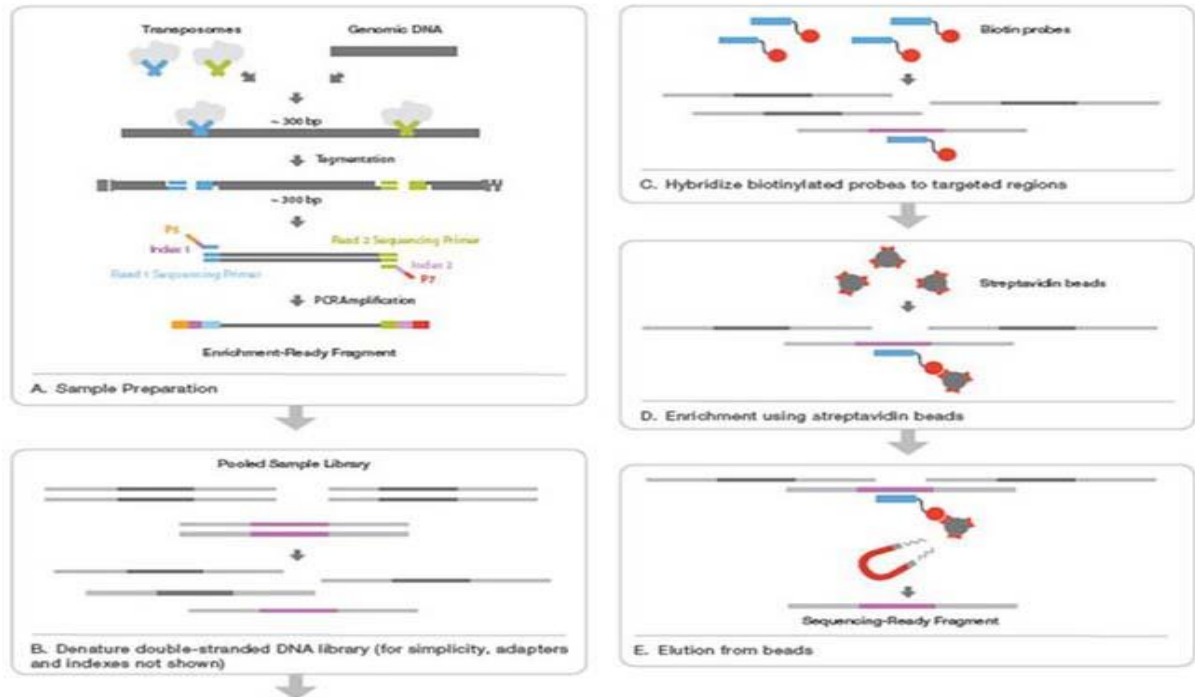


Figura 13. Representación esquemática de la preparación de las librerías mediante el exoma clínico, utilizando la tecnología de TruSight One (Illumina). Imagen adaptada del protocolo de Illumina.

3.2.5. Priorización y clasificación de las variantes

Se priorizaron variantes en heterocigosis en genes previamente asociados con adRP y cuando no se encontraron variantes candidatas, se consideraron variantes asociadas con otros patrones de herencia (ar o xl) u otras presentaciones clínicas distintas a la RP.

Además, se siguió un protocolo de priorización de variantes potencialmente patogénicas teniendo en cuenta:

- i) localización en exones y regiones canónicas de *splicing*, que presentaron un *score* de calidad >100 en su secuenciación.
- ii) frecuencia alélica descrita en gnomAD (<http://gnomad.broadinstitute.org/>). En función del tipo de herencia estudiada, dominante o recesiva, se consideraron variantes raras con una frecuencia inferior al 0,1% o 1%, respectivamente.
- iii) identificación previa como causante de DR en distintas bases de datos (HGMD, LOVD, ClinVar).

iv) tipo de consecuencia de la variante sobre la proteína. Se consideraron como patogénicas:

- variantes *nonsense* y *frameshift* que causan un truncamiento de la proteínas (*LOF*, *loss-of function*).
- las variantes de *splicing*, cuyo análisis *in silico* predecía una posible alteración en el procesamiento del ARNm.
- variantes *missense* altamente conservadas evolutivamente, que fueron predichas como patogénicas en al menos 2 de los 4 predictores de patogenicidad *in silico* utilizados.

Durante la valoración de la patogenicidad de variantes nuevas de *splicing* o *missense*, se utilizaron al menos los predictores *in silico* implementados en el programa comercial *Alamut* (Interactive Biosoftware, Rouen, France): *Human Splicing Finder*, *MaxEntScan*, *NNSPLICE*, *GeneSplice* y *ESEFinder* para las variantes de *splicing* y *SIFT*, *Polyphen*, *MutationTaster* y *Align GVGD* para las variantes tipo *missense*.

Todas las variantes identificadas en el análisis de NGS, clasificadas como VUS (*variant of unknown significance*), patogénicas o probablemente patogénicas, se confirmaron mediante secuenciación de Sanger y fueron segregadas en las familias utilizando muestras disponibles de familiares afectos y sanos.

La clasificación de la causalidad de las variantes se estableció tal como se detalla a continuación (Figura 14), siguiendo las recomendaciones del *American College of Medical Genetics and Genomics* (Richards *et al.*, 2015) y de la Sociedad Europea de Genética Humana (Matthijs *et al.*, 2016):

i) **probablemente no causales**: variantes raras en las que la segregación en la familia las excluye como responsables. Incluye las variantes que se clasifican habitualmente como tipo 1 (benigna) y 2 (probablemente benigna).

ii) **variantes de significado incierto** (VUS): variantes nuevas tipo *missense* e *in-frame* patogénicas según predictores *in-silico*, con segregación no informativa en familias o no disponible.

iii) **variantes probablemente causales**: variantes nuevas tipo *missense* e *in-frame* con predicción patogénica según predictores *in-silico*, y con segregación con la RP en familias informativas. Corresponden a las variantes tipo 4 (variante probablemente patogénica), según la nomenclatura internacional.

iv) **variantes causales**: variantes patogénicas previamente asociadas a RD en la literatura y/o bases de datos; nuevas variantes LOF (*frameshift*, *nonsense*, *splicing* en sitios canónicos con

predictores patogénicos). Corresponden a las variantes tipo 5 (variante patogénica), según la nomenclatura internacional.

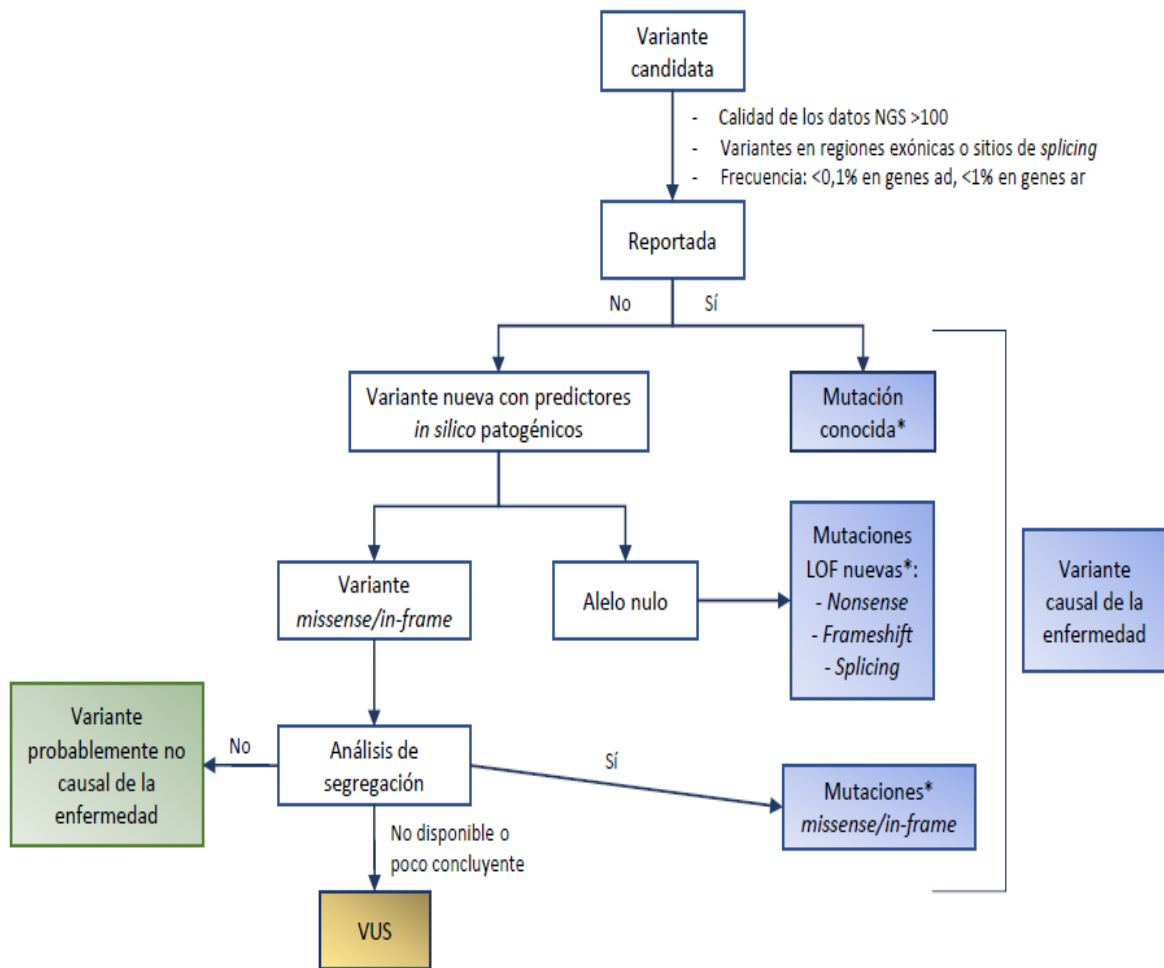


Figura 14. Algoritmo llevado a cabo para la clasificación de las variantes. *Análisis de segregación cuando fue posible. ad: autosómico dominante; ar: autosómico recesivo. Imagen adaptada de Fernandez-San Jose *et al.*, 2015a.

3.3. Secuenciación de Sanger

La amplificación del ADN por PCR se llevó a cabo usando el kit comercial *FastStart Taq DNA Polymerase* (Roche Life Science), en los termocicladores *Veriti Thermal Cyclers* o *GeneAmp PCR System 2700* (Applied Biosystems). Los productos de PCR obtenidos se comprobaron mediante migración electroforética en gel de agarosa al 2% en solución TBE 1X. Los productos de PCR fueron procesados por el método de purificación enzimática *ExoSAP-IT* (USB Corporation Affymetrix).

Posteriormente, se realizó la reacción de secuenciación empleando el kit comercial *Big Dye Terminator v3.1 Cycle Sequencing* (Applied Biosystems). La purificación del producto de secuenciación se realizó utilizando el kit comercial *Montage SEQ96 Sequencing Reaction Cleanup*

(Merck Millipore), basado en placas de filtración. Una vez purificado el producto se introdujo en el secuenciador automático ABI3130xl (Applied Biosystems, CA, USA). El análisis de las secuencias se realizó mediante los programas informáticos *Staden Package* y/o *Novosnp 3.0.1*.

3.4. Análisis de CNVs

El análisis de CNVs de los genes asociados a adRP y su posterior validación, se realizó mediante las distintas aproximaciones que se describen a continuación:

3.4.1. MLPA

En este trabajo se utilizaron kits comerciales de MLPA (MRC-Holland, Amsterdam, Netherlands) para distintos genes asociados a adRP:

i) kit P235 con sondas para los genes *RHO*, *IMPDH1*, *PRPF31* y *RP1*, utilizado en 103 familias no caracterizadas después del abordaje clásico o el panel de genes customizado usando la tecnología HaloPlex.

ii) kit P221 con sondas para el gen *CRX*, entre otros genes, que fue utilizado durante la validación de CNVs en este gen identificadas con otras aproximaciones durante este trabajo.

El MLPA se realizó siguiendo las instrucciones del fabricante. Los fragmentos amplificados fueron separados por electroforesis capilar en el secuenciador automático ABI 3130xl (Applied Biosystems) y analizados usando los programas GeneMapper (Applied Biosystems) y Coffalyser (MRC-Holland). Los datos se interpretaron a través del ratio de la señal de cada sonda entre el ADN del paciente y el control. Un ratio de 1 indicó la presencia de dos alelos (diploide normal), mientras que un ratio por debajo de 0,6 o por encima de 1,4 sugirió una delección o una duplicación de la secuencia diana, respectivamente. En los casos positivos, el MLPA se extendió a otros miembros de la familia disponibles para completar el análisis de segregación.

3.4.2. Análisis de CNVs con datos de NGS

En las 68 muestras con adRP previamente analizadas mediante NGS con la tecnología HaloPlex, los archivos Fastq fueron reanalizados utilizando una *pipeline* implementada en nuestro laboratorio recientemente. Para ello, se utilizó el programa CoNVading (*Copy Number Variation Detection In Next-generation sequencing Gene panels*) (Johansson *et al.*, 2016), con los parámetros por defecto del programa, que permite detectar pequeñas CNVs a nivel de exones, usando datos de NGS con alta profundidad de las regiones capturadas.

Así, se calculó la profundidad de lectura de cada región capturada de una muestra seleccionada como problema entre la media de todas las demás muestras consideradas como controles en la

misma carrera. Un ratio de 1 indicó la presencia de dos alelos (diploide normal), mientras que un ratio por debajo de 0,6 o por encima de 1,4 sugirió una delección o una duplicación de la región diana, respectivamente.

Para validar esta *pipeline* se usaron como controles positivos un total de 8 muestras de NGS con CNVs conocidas, cuya librería se preparó mediante HaloPlex, incluyendo una familia adRP con una duplicación en el gen *PRPF31* previamente detectada por MLPA. Todas estas CNVs fueron correctamente detectadas mediante este método.

Las muestras en las que se detectaron nuevas CNVs mediante esta aproximación fueron validadas con una segunda, bien MLPA o arrays de hibridación genómica comparada (aCGH), en función del tipo de reordenamiento, tamaño y cobertura del gen a analizar en plataformas comerciales.

3.4.3. Arrays de hibridación genómica comparada

En este trabajo se utilizó la técnica de aCGH para delimitar los reordenamientos genómicos previamente identificados por otras técnicas. Para ello, se utilizaron arrays SurePrint G3 de la plataforma Agilent Technologies (Santa Clara, USA) siguiendo los protocolos recomendados para el marcaje, hibridación, escaneado y análisis de los arrays. En resumen, 500 ng de ADN procedente de los individuos en estudio y ADN de referencia del mismo sexo (Promega, Madison, WI, USA) fueron digeridos con RsaI y AluI durante 2h a 37°C. Tras la inactivación de las enzimas a 65°C durante 20 minutos, las muestras digeridas fueron marcadas utilizando el kit *Genomic DNA Enzymatic Labelling Kit* (Agilent) utilizando los fluoróforos Cy5-dUTP y Cy3-dUTP para el ADN de los pacientes y la referencia, respectivamente. Posteriormente, las muestras marcadas fueron purificadas utilizando las columnas Ym-30 (Merck Millipore Corporation, Darmstadt, Germany) e hibridadas en los cristales a 65°C durante 24h. Tras su lavado, los cristales fueron escaneados a 3 µm de resolución utilizando el programa *Feature Extraction* (Agilent).

El análisis y visualización de los datos se realizó con el programa *Agilent CytoGenomics v.2.7* usando el método de análisis *Aberration Detection 2* (ADM-2) para calcular los *log2-ratio* de las intensidades de fluorescencia relativas del paciente (rojo, Cy5) *versus* al ADN de referencia (verde-Cy3), para cada una de las sondas. Los ratios cercanos a -1 indicaron la presencia de una delección de 1 copia (heterocigosis), los ratios cercanos a 0 presentaron un patrón normal, mientras que los ratios cercanos a +0,5 indicaron amplificaciones heterocigotas. La presencia de una CNV significativa vino determinada por la presencia de al menos 3 sondas consecutivas delecionadas o amplificadas.

En este trabajo se utilizaron varios tipos de aCGH diferentes, en función de la cobertura del gen a analizar:

i) **array de 400k** (*PostNatal Research*, SurePrint G3 CGH+SNP 2X400k), que contiene 292.097 sondas CGH, con un espaciamiento medio de las sondas de 3,7 kb y 119.091 sondas para análisis de SNPs. Este array fue utilizado específicamente para analizar 4 individuos portadores de distintas deleciones y/o duplicaciones en el gen *PRPF31*.

ii) **array de 60K sondas**, que fue utilizado para analizar 2 individuos portadores de deleciones en el gen *CRX*.

3.4.4. Long-Range PCR

Con el fin de analizar si una amplificación de 5,1 kb en el gen *PRPF31*, identificada en este trabajo, podría tratarse de una duplicación en tándem se utilizó una estrategia de *long-range PCR*.

Se diseñaron cebadores específicos en las regiones adyacentes a los puntos de corte de la amplificación (Exón 1F: ACTTGTCCTCATCAGGCTCAG; Exón 6-7R: TCCCAAGGTCACAGTGTCAG), previamente identificados mediante aCGH, con el fin de poder discriminar entre el alelo *wild-type*, con un tamaño de 8,6 kb, de una posible duplicación en tándem, que tendría un tamaño esperado de 13,7 kb. Para ello, se utilizó el kit comercial *Expand Long Range dNTPack* (Roche Life Science, Basel, Switzerland), siguiendo las instrucciones del fabricante (Tabla 3). Los amplicones de la PCR se comprobaron en un gel de agarosa al 1%, usando como marcador Lambda DNA/HindIII.

Tabla 3. Condiciones de la *long-Range PCR* para la amplificación de la duplicación en el gen *PRPF31*. Volumen final de 50 µl. *Cebadores: 1,5 µl del *forward* + 1,5 µl del *reverse*. *0,20'' extra por ciclo.

								95°	2'		
								95°	10''		
								60°	15''	10 CICLOS	
								72°	14'		
Volúmen final	Agua	DMSO	Tampón con Mg ²⁺ 5x	dNTPs 500mM	Cebadores* (10µM)	ADN	Enzima 5U/µl				
50 µl	30,8 µl	2 µl	10 µl	2,5 µl	3 µl	1,5 µl	0,7 µl	95°	10''		
								60°	15''	25 CICLOS	
								68°	*14'		
								68°	7'		
								4°	∞		

3.5. Análisis de haplotipos

En varias familias de este trabajo, se realizó un análisis de haplotipos para analizar la posible existencia de un efecto fundador en los genes *PRPF31* y *CRX*. Para ello, se utilizaron los marcadores tipo microsatélite (*short tandem repeat*, STR) polimórficos con elevada heterocigosidad, que fueron seleccionados utilizando la base de datos GeneLoc. En algunos casos, se utilizaron también marcadores bialélicos tipo SNPs con MAF elevada, seleccionados de la base de datos dbSNP.

Los microsatélites se analizaron mediante análisis de fragmentos. Primero, fueron amplificados por separado utilizando parejas de cebadores en donde el cebador *forward* estaba marcado con un fluoróforo para permitir su posterior detección por fluorescencia de los productos amplificados mediante electroforesis capilar en el equipo ABI3130xl (Applied Biosystems, CA, USA). El análisis de fragmentos se realizó con el programa *GeneMapper* v3.5 (Applied Biosystems). Los marcadores tipo SNPs fueron genotipados mediante secuenciación Sanger, utilizando cebadores específicos diseñados mediante Primer3. En la construcción de los haplotipos se utilizó el programa *Cyrillic* (Cyrillic Software, Wallingford, UK).

En el análisis de haplotipos para el gen *PRPF31*, se utilizaron cinco marcadores microsatélite con gran heterocigosidad (D19S921, D19S572, D19S927, D19S418 y D19S605) flanqueando 1,9 Mb alrededor del gen en 6 familias para estudiar un posible efecto fundador de una mutación (en 4 familias) o establecer la fase de los alelos del minisatélite MSR1 adyacente al promotor de *PRPF31* (en 2 familias). Además, seis polimorfismos de un solo nucleótido (SNP: *single-nucleotide polymorphisms*) (rs57960425, rs2303557, rs1058572, rs11556769, rs76781318, y rs655240) fueron genotipados en 2 familias con la misma mutación causal (Tabla 4).

Tabla 4. Microsatélites y SNPs del gen *PRPF31*.

	Marcador/ SNPs	Inicio (Mb)	Amplicón (pb)	Fluorocromo	
<i>PRPF31</i> _chr19: 54115457-54131709	D19S921	53,3	234	FAM	1,9 Mb
	D19S572	53,6	134	FAM	
	D19S927	53,8	141	HEX	
	rs57960425	c. 528-78_528-75delCACA			
	rs2303557	c.528-31C>T			
	c.541G>T	p.Glu181*			
	rs1058572	c.564G>A			
	rs11556769	c.735C>T			
	rs76781318	c.855+40G>A			
	rs655240	c.1148-9T>C			
	D19S418	55,0	100	HEX	
	D19S605	55,2	117	FAM	

En otras 2 familias aparentemente portadoras de la misma delección afectando varios exones del gen **CRX**, se construyeron haplotipos utilizando 3 STRs flanqueando 1,5Mb alrededor del gen **CRX** (D19S606, D19S596 y D19S879) y un marcador microsatélite intragénico (D19S902), con el fin de determinar un posible ancestro común (Tabla 5).

Tabla 5. Microsatélites del gen **CRX**.

	Marcador	Inicio (Mb)	Amplicón (pb)	Fluorocromo	
CRX_chr19: 48325097-48346586	D19S606	48,0	178	NED	1,5 Mb
	D19S902	48,3	214	FAM	
	Del Ex. 3-4				
	D19S596	49,3	213	NED	
	D19S879	49,5	253	VIC	

3.6. Análisis de expresión génica

El ARN total de 3 individuos con reordenamientos en el gen **PRPF31** y 9 controles se transcribió usando cebadores aleatorios (*random primers*) con el kit *High Capacity cDNA Reverse Transcription* (Applied Biosystems) en un volumen final de 20 µL. Se realizó una PCR cuantitativa (qPCR) utilizando el intercalante *SYBR Green PCR Master Mix kit* (Roche Life Science) y cebadores específicos (Exón 7-8F: CCAAGATCATGGGTGTGG; Exón 8R: GGGGCAGCACTGAGGTAG) en el sistema *LightCycler 480 Real-Time PCR* (Roche Life Science). El gen gliceraldehido-3-fosfato deshidrogenasa (GAPDH) se utilizó como control génico endógeno. Todas las reacciones se realizaron por triplicado y en dos análisis independientes. La intensidad de la fluorescencia se analizó usando el programa *LightCycler 480 gene scanning* (Roche Life Science) y el método de $\Delta\Delta C_t$ se empleó para calcular las diferencias en las abundancias relativas del ARN mensajero del gen **PRPF31** respecto a las muestras controles. El test de Mann-Whitney se utilizó para determinar la significancia estadística ($p\text{-valor} \leq 0,05$) de las diferencias en la abundancia relativa de **PRPF31** entre los dos grupos.

3.7. Análisis de CNVs del minisatélite MSR1 adyacente al promotor de PRPF31

Se analizó el número de copias del minisatélite MSR1, localizado a unas 200 pb aguas arriba del promotor de **PRPF31**, en 35 pacientes sintomáticos y en 5 pacientes asintomáticos de 15 familias con mutaciones en el gen **PRPF31**, así como en una cohorte de 123 controles. El genotipado de este minisatélite se realizó mediante un análisis de fragmentos utilizando cebadores específicos previamente diseñados (Rose *et al.*, 2016) (Tabla 6). Los alelos de 3 copias y 4 copias, que

presentan un tamaño de 462 bp y 500 bp, respectivamente fueron discriminados mediante electroforesis capilar en un secuenciador automático ABI 3130xl (Applied Biosystems) y los resultados se analizaron usando el programa GeneMapper (Applied Biosystems) (Figura 15). Se utilizó el test de proporciones o el test exacto de Fisher, cuando la frecuencia esperada fue menor de 5, para valorar si existían diferencias con significancia estadística en las frecuencias de los alelos de MSR1 entre los distintos grupos (sintomáticos, asintomáticos y controles).

Tabla 6. Condiciones de la PCR para la amplificación de MSR1. Secuencia de los cebadores usados para la PCR. Volumen final de 15 µl. *Cebadores: 1 µl del *forward* + 1 µl del *reverse*.

Primer forward: 5´FAM-GTTAGGGGTTTGGACTGC								94°	12´	10 CICLOS
Primer reverse: GATGTGGCCACCAAATAG								94°	30´´	
								55°	45´´	
								72°	1´	
Volúmen final	Agua	DMSO	Tampón 10x	dNTPs	Cebadores* (10µM)	ADN	Enzima			
15 µl	5,3 µl	2 µl	1,5 µl	3 µl	2 µl	1 µl	0,2 µl	89°	30´´	20 CICLOS
								55°	45´´	
								72°	1´	
								72°	30´	
								4°	∞	

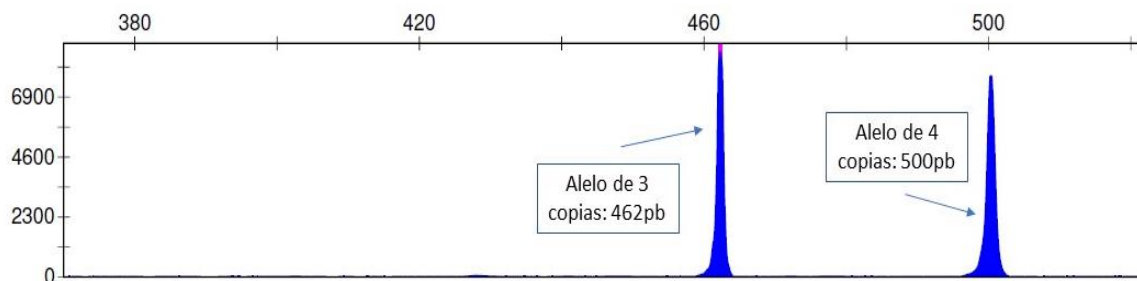


Figura 15. Imagen de un individuo heterocigoto 3/4 para el minisatélite MSR1. Obtenida del programa GeneMapper.

3.8. Bases de datos y herramientas bioinformáticas

Las bases de datos y herramientas bioinformáticas utilizadas a lo largo del estudio son:

Bases de datos:

- Ensembl: <http://www.ensembl.org>
- HGMD: <http://www.hgmd.cf.ac.uk>
- RetNet: <https://sph.uth.edu/retnet>

- UCSC: <https://genome.ucsc.edu/>
- gnomAD: <http://gnomad.broadinstitute.org/>
- ClinVar: <https://www.ncbi.nlm.nih.gov/clinvar/>
- GeneLoc: <https://genecards.weizmann.ac.il/geneloc/>
- dbSNP: <https://www.ncbi.nlm.nih.gov/projects/SNP/>

Herramientas para análisis bioinformático de NGS:

- BWA (*Burrows-Wheeler Aligner*)-MEM: <https://bio-bwa.sourceforge.net/>
- GATK (*Genome Analysis Toolkit*): <https://software.broadinstitute.org/gatk/>
- ENSEMBL Variant Effect Predictor: <https://www.ensembl.org/vep>
- ANNOVAR: <https://annovar.openbioinformatics.org/>
- CoNVadIng: <https://github.com/molgenis/CoNVaDING>
- DNANexus <https://www.dnanexus.com/>

Predictores *in-silico*:

Splicing:

- Human Splicing Finder: <https://www.umd.be/HSF3/>
- MaxEntScan: https://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html
- NNSPLICE: <https://omictools.com/nnsplICE-tool>
- GeneSplicer: https://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml
- ESEFinder: <https://rulai.cshl.edu/>
- Spidex: <http://tools.genes.toronto.edu/>

Missense:

- SIFT: <https://sift.jcvi.org/>
- Polyphen: <https://genetics.bwh.harvard.edu/pph2/>
- MutationTaster: <https://www.mutationtaster.org/>
- Align GVGD: <https://www.mutationtaster.org/>

- M-CAP: <https://bejerano.stanford.edu/mcap/>
- CADD: <https://cadd.gs.washington.edu/>

Resultados

Durante el desarrollo de esta Tesis Doctoral, se ha estudiado una cohorte de 258 familias españolas con sospecha de adRP, recogida entre los años 1991 y 2017 en la consulta de Genética del Hospital Universitario Fundación Jiménez Díaz. El presente trabajo supone una continuación y ampliación del trabajo de varios proyectos de investigación y tesis doctorales realizadas durante este periodo, incluida la más reciente, la Tesis doctoral de Patricia Fernández San José (Fernández-San-José, 2017), donde se valoró la utilidad de la NGS en la caracterización genética de la adRP en un total de 65 familias. En el presente estudio, se ha querido abordar un análisis completo de nuestra cohorte con esta tecnología de NGS y además valorar la implicación de las CNVs en la adRP. Todo ello con el fin de obtener una mejor caracterización molecular de esta patología y una estimación de su espectro mutacional en población española. Para ello, a lo largo de esta tesis doctoral se ha producido:

- i) Reclutamiento de 58 nuevas familias con sospecha de adRP.
- ii) Estudio de CNVs mediante distintas aproximaciones.
- iii) Evaluación del rendimiento de una nueva *pipeline* para el estudio simultáneo de SNVs y CNVs.
- iv) Evaluación de una nueva estrategia de NGS basada en la tecnología MIPs.
- v) Evaluación del rendimiento diagnóstico de la estrategia de exoma clínico para el estudio genético de la adRP.
- vi) Estudio exhaustivo de la implicación del gen *PRPF31*, el segundo gen más prevalente en familias españolas con adRP.

Todo ello nos ha permitido además hacer una evaluación de la eficacia diagnóstica de la NGS en las formas dominantes de RP en una de las mayores cohortes de pacientes reclutadas a nivel mundial. A lo largo del presente trabajo, los resultados se expondrán en 4 bloques:

1. Cribado mutacional de 258 familias españolas con adRP mediante distintas tecnologías.
2. Análisis de variaciones en el número de copias (CNVs).
3. Tasa global de detección de mutaciones en 258 familias con adRP en población española.
4. Estudio del gen *PRPF31* en familias españolas con adRP.

4.1. Cribado mutacional de 258 familias españolas con adRP, mediante distintas tecnologías moleculares

Las familias han sido analizadas utilizando diferentes aproximaciones moleculares a lo largo del tiempo (Figura 16). En resumen, el estudio de las 258 familias con adRP se realizó siguiendo en primer lugar uno de los dos abordajes metodológicos, dependiendo de la época en que se reclutó la familia:

- i. **Abordaje clásico** en 226 familias, caracterizándose un total de 87 familias, de las cuales 46 familias mediante la detección directa o indirecta de los genes más prevalentes y 41 familias mediante la detección de mutaciones conocidas a través del chip de genotipado de adRP.
- ii. **Abordaje mediante exoma clínico**, utilizando un panel virtual en 32 familias *naive* (no cribadas anteriormente) que acudieron a consulta en una etapa posterior, resolviéndose 21 de ellas.

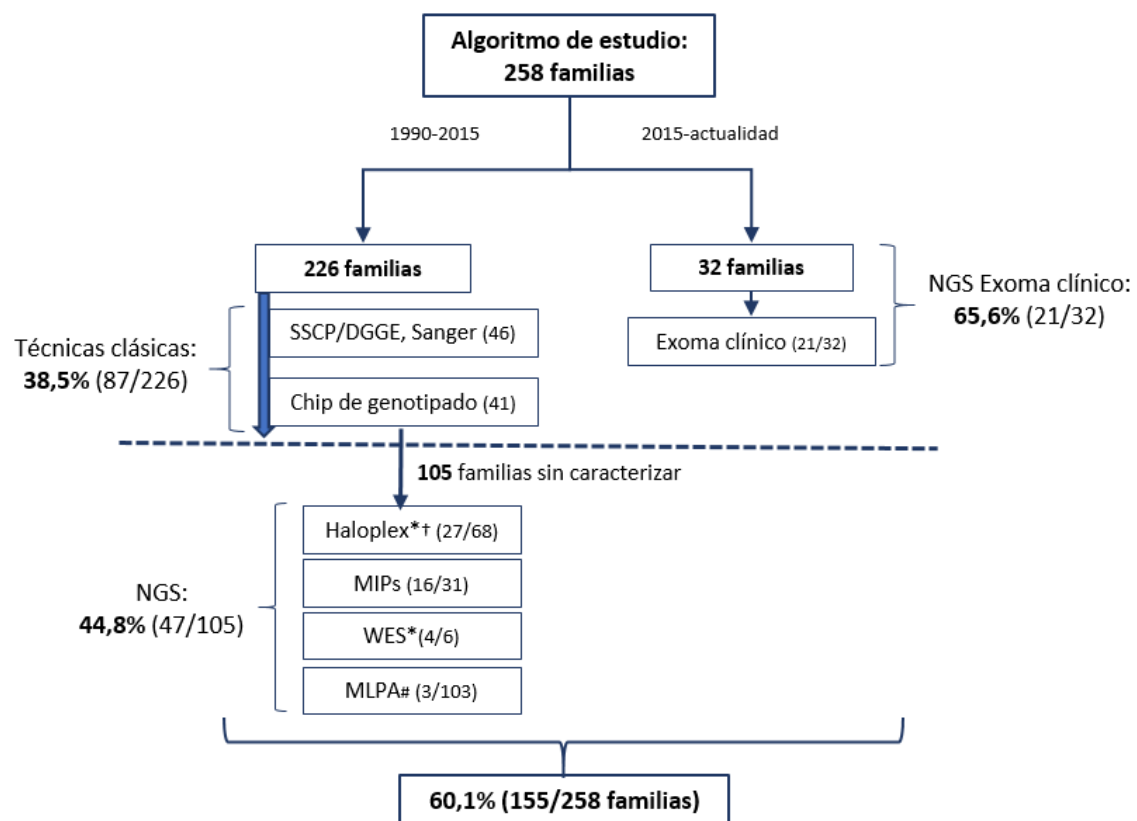


Figura 16. Algoritmo usado para estudiar a las 258 familias con adRP. Se muestra el porcentaje de caracterización obtenido mediante el uso de las técnicas clásicas y la NGS (mediante exoma clínico) como primer abordaje, la NGS a las familias sin caracterizar tras las técnicas clásicas y la tasa diagnóstica global en la cohorte. *NGS para análisis de SNV y CNV. † Reanálisis bioinformático. #Una familia con dos reordenamientos distintos.

En una segunda etapa, 105 familias sin caracterizar tras el cribado molecular clásico fueron analizadas mediante NGS, utilizando un panel de genes customizado (HaloPlex o MIPs), o WES, y/o MLPA (apartado 4.2.1), lo que permitió caracterizar a 47 de ellas.

Adicionalmente, y como tercer paso, en el caso de las familias adRP estudiadas mediante el panel de HaloPlex (68 familias) se realizó un reanálisis bioinformático, mediante una *pipeline* diseñada *in house*, para el estudio de SNVs y de CNVs. Se detectaron las mismas variantes causales identificadas en el análisis inicial y además se encontraron otras 8 nuevas variantes en 7 familias, incluyendo 5 SNVs (Tabla 7) y 3 CNVs, estas últimas serán descritas con más detalle en el siguiente apartado (4.2), lo que permitió caracterizar a un 10,3% (7/68) de las familias estudiadas mediante esta aproximación.

Tabla 7. SNVs causales tras el reanálisis de un panel de NGS. Las 68 muestras analizadas con el panel de HaloPlex fueron reanalizadas mediante una *pipeline* diseñada *in house*, encontrándose 5 SNVs nuevas en 4 familias.

Familia	Gen	Exón	Cambio nucleotídico	Cambio proteína	Referencia
Genes dominantes					
RP-0187	<i>PRPF31</i>	4	c.322+4_322+7del	<i>Splicing</i>	(Q. Zhang <i>et al.</i> , 2016)
RP-1118	<i>TOPORS</i>	3	c.2518_2519del	p.Ser840*	Presente estudio
Genes recesivos					
RP-0038	<i>CNGA1</i>	5	c.94C>T	p.Arg32*	(Paloma <i>et al.</i> , 2002)
		6	c.131del	p.Glu44Glyfs*49	Presente estudio
RP-1217	<i>EYS</i>	IVS28	c.5928-2A>G	<i>Splicing</i>	(Gonzalez-del Pozo <i>et al.</i> , 2011)

4.2. Análisis de variaciones en el número de copia (CNV)

En este estudio, se quiso valorar la posible presencia de CNVs en los genes asociados a adRP utilizando distintas estrategias de análisis. Para ello, se analizaron algunos de los genes más prevalentes mediante MLPA. Posteriormente, se realizó un análisis bioinformático específico basado en la profundidad de cobertura de secuencias a partir de los datos de NGS. En total, identificamos 8 CNVs diferentes en 3 genes asociados a adRP en 7 familias distintas (Tabla 8).

Tabla 8. Grandes reordenamientos genómicos encontrados en nuestra cohorte. Los reordenamientos se encontraron en las posiciones 19q13.42, 19q13.32 y 3q22.1, afectando a los genes *PRPF31*, *CRX* y *RHO*, respectivamente. En negrita: variantes descritas en nuestra población por primera vez.

Familia	Tipo de reordenamiento	Genes implicados	Técnica	Validación	Coordenadas genómicas (hg19)	Tamaño mínimo(pb)	Tamaño máximo(pb)	Referencia
RP-0777	Deleción	<i>OSCAR</i> (E1 al E2), <i>NDUFA3</i> , <i>TFPT</i> y <i>PRPF31</i>	NGS (WES)	aCGH	chr19:54,602,946-54,635,178	32,232	111,184	(Almoguera <i>et al.</i> , 2015)
RP-0932	Deleción	<i>OSCAR</i> (E1 al E2), <i>NDUFA3</i> , <i>TFPT</i> y <i>PRPF31</i> (E1 al E13)	MLPA	aCGH	chr19:54,602,946-54,632,693	29,747	87,435	Presente estudio
RP-0932	Duplicación	<i>PRPF31</i> (E2 al E5)	MLPA	aCGH	chr19:54,621,606-54,626,745	5,139	7,832	Presente estudio
RP-2426	Deleción	<i>PRPF31</i> (E1)	MLPA	aCGH	chr19:54,618,828-54,619,055	227	18,609	Presente estudio
RP-0076	Deleción	<i>PRPF31</i> (E9)	NGS (Haloplex)	MLPA	chr19:54,629,903-54,629,992	89	3,435	Presente estudio
RP-1092	Deleción	<i>CRX</i> (E3 al E4)	NGS (Haloplex)	MLPA+aCGH	chr19:48,339,216-48,345,529	6,313	12,007	(Bravo-Gil <i>et al.</i> , 2016)
RP-1192	Deleción	<i>CRX</i> (E3 al E4)	NGS (Haloplex)	MLPA+aCGH	chr19: 48,339,345-48,347,355	8,010	11,905	(Bravo-Gil <i>et al.</i> , 2016)
RP-1383	Deleción	<i>RHO</i> (E5)	MLPA	Sanger	chr3:129,251,762-129,252,589	827	827	(de Sousa Dias <i>et al.</i> , 2015)

4.2.1. Análisis de los reordenamientos mediante MLPA

En 103 *probandos* no caracterizados tras realizar técnicas clásicas de cribado o un panel customizado de genes (HaloPlex), se continuó el estudio mediante análisis de CNV en los genes *RHO*, *PRPF31*, *RP1* e *IMPDH1*, con la realización del MLPA (P235), para detectar posibles deleciones o duplicaciones en estas familias. De esta manera se encontraron 3 reordenamientos diferentes: una deleción del exón 1 en el gen *PRPF31* en la familia RP-2426, una deleción del exón 5 en *RHO* (de Sousa Dias *et al.*, 2015) en la familia (RP-1383) y una duplicación de los exones 2 al 5 afectando a *PRPF31* en la familia RP-0932. El análisis de segregación mediante MLPA en esta última familia permitió detectar que se encontraba segregando además de la duplicación en el gen *PRPF31* en una rama de la familia, una deleción de los exones 1 al 13 en el mismo gen en otra rama de la familia (Figura 17). En resumen, el MLPA nos permitió detectar 4 reordenamientos distintos en 3 familias (Tabla 8).

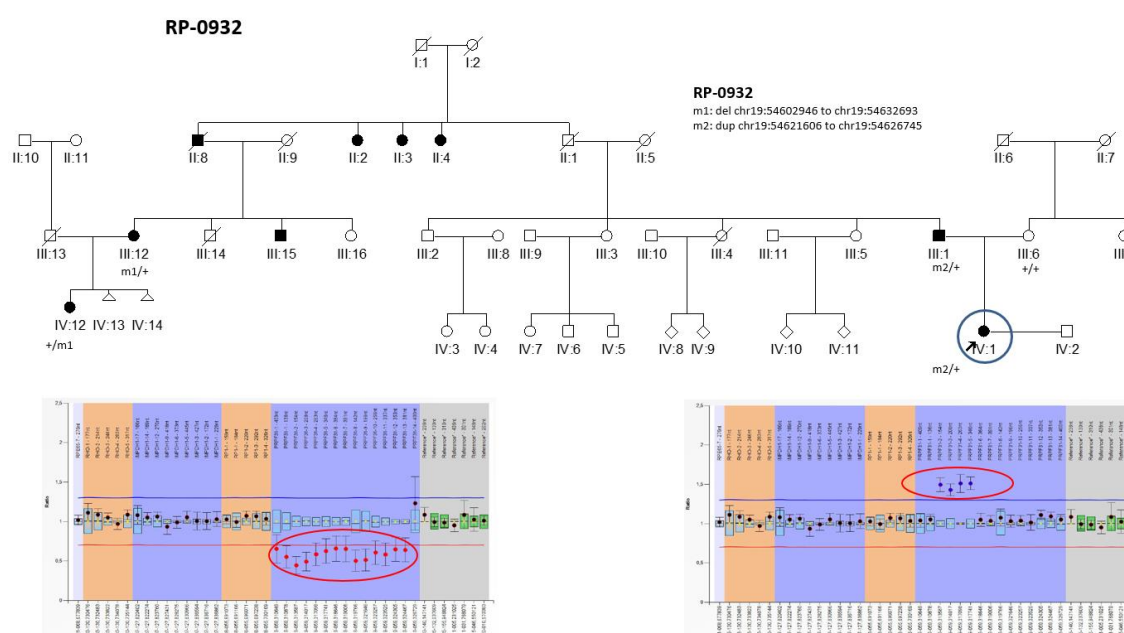


Figura 17. Árbol y segregación en la familia RP-0932 con dos reordenamientos en *PRPF31*. El análisis de segregación mediante MLPA puso de manifiesto la presencia de una duplicación de los exones 2 al 5 en el gen *PRPF31* en la *probandus* (IV:1) y su padre afecto (III:1), y una deleción de los exones 1 al 13 en el gen *PRPF31* en su prima paterna (IV:12) y su tía paterna (III:12).

4.2.2. Análisis de los reordenamientos mediante datos de NGS

Esta última aproximación fue realizada en 68 casos que habían sido analizados previamente con un panel de genes personalizado con la aproximación de HaloPlex, secuenciados a una alta cobertura de secuencia de >450X. El análisis de cobertura de datos de NGS confirmó la presencia de las 3 deleciones encontradas previamente mediante MLPA (duplicación de los exones 2 al 5 en *PRPF31*, deleción del exón 1 de *PRPF31* y deleción del exón 5 en *RHO*) además, permitió detectar nuevas deleciones de 1 copia afectando múltiples exones o incluso a tan solo un exón, en los genes *CRX* y *PRPF31* (Figura 18). En concreto, se identificó una deleción del exón 9 en el gen *PRPF31* en la familia RP-0076 y una deleción de los exones 3 y 4 del gen *CRX* en 2 familias de adRP (RP-1092 y RP-1192), que fueron posteriormente confirmadas las deleciones y segregadas en las familias mediante análisis de MLPA. En total, junto a la gran deleción en el gen *PRPF31* en la familia RP-0777 encontrada mediante WES (Almoguera *et al.*, 2015), se encontraron 4 CNVs en 4 familias.

En el caso de la deleción aparentemente idéntica encontrada en el gen *CRX* en 2 familias, se realizó un análisis de haplotipos en la región cercana a la deleción, permitiendo descartar un posible origen común ancestral para esta CNV y por tanto, un posible efecto fundador (Figura 18:C).

4.2.3. Validación de los reordenamientos en *PRPF31* y *CRX* mediante aCGH

En todos aquellos reordenamientos que incluían la deleción del primer o último exón del gen, se realizó un aCGH por la posibilidad de que las deleciones detectadas mediante MLPA fueran mayores e incluyeran genes adyacentes, permitiéndonos así delimitar en mayor grado los reordenamientos. Además, la duplicación hallada en la familia RP-0932 en *PRPF31* también se validó mediante esta tecnología.

Así en 3 familias portadoras de deleciones en el gen *PRPF31*, los puntos de corte se definieron usando un aCGH comercial de sondas de 400 K que cubren el locus del gen con 34 sondas. Para la familia RP-0932, el tamaño de la deleción fue de al menos 29,7 kb, abarcando no solo parcialmente *PRPF31* sino también tres genes contiguos aguas arriba (*TFPT*, *NDUFA3* y parte del gen *OSCAR*), como se detalla en la Figura 19 y Tabla 8. Además, el análisis de aCGH nos permitió delimitar la duplicación, identificada en la otra rama familiar, con un tamaño mínimo de 5,1 kb que involucraba solo parte de *PRPF31*.

En la familia RP-2426, el análisis de aCGH confirmó una pequeña deleción de un tamaño mínimo de 227 pb, que abarcaba solo el exón 1 del gen *PRPF31*, como se había detectado previamente usando el MLPA (Figura 19, Tabla 8).

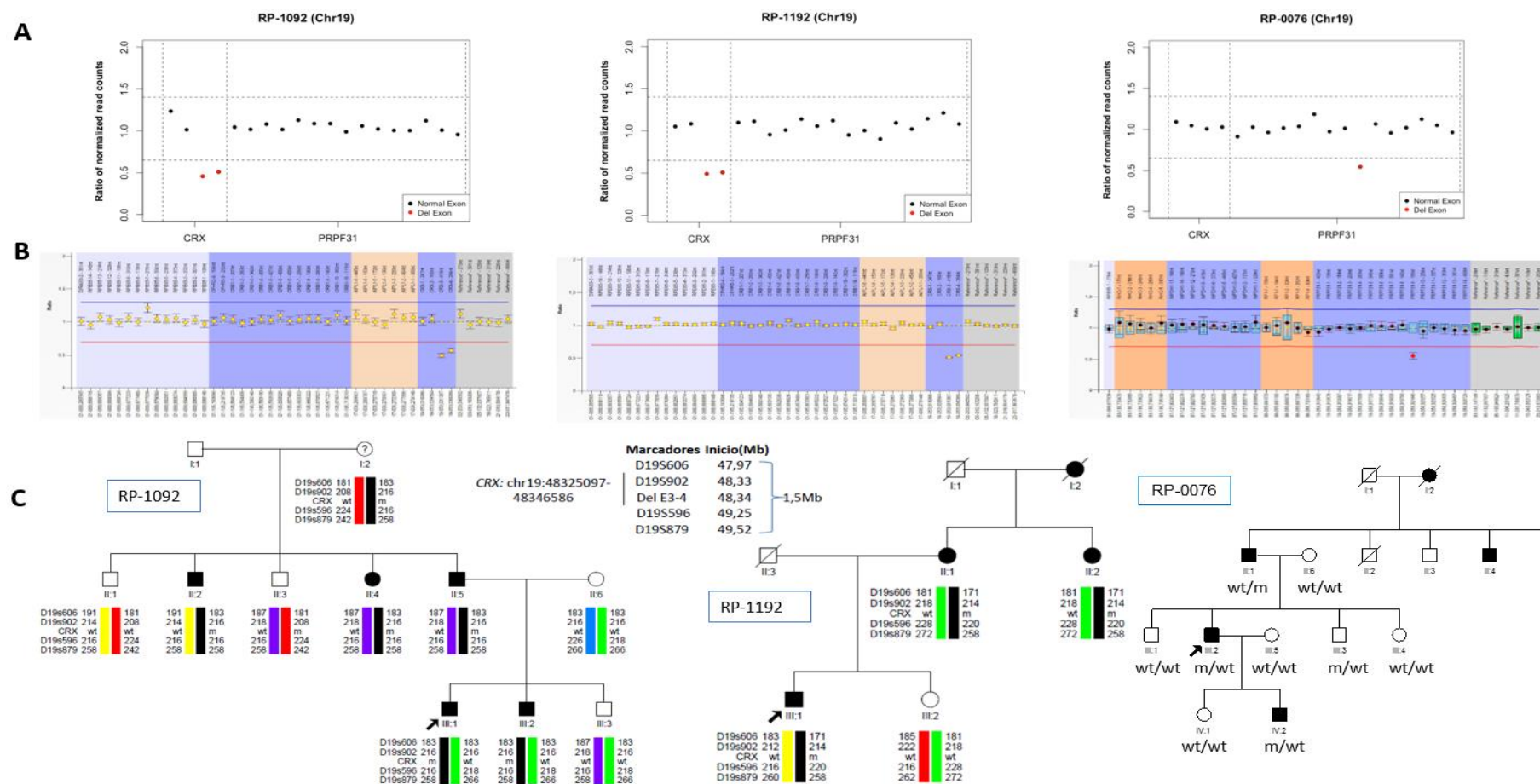


Figura 18. Resultados de los reordenamientos mediante el análisis de la profundidad de lectura de los datos de NGS. **A.** Representación esquemática de las 3 deleciones encontradas en los genes *CRX* y *PRPF31* (cromosoma 19) en las familias RP-1092, RP-1192 y RP-0076. Cada exón está representado por un círculo. Los exones normales se representan en negro, mientras que los deletados en rojo. **B.** Validación de las deleciones mediante MLPA. **C.** Análisis de haplotipos usando tres marcadores microsatélite (D19S606, D19S596 y D19S879) flanqueando 1,5 Mb alrededor del gen *CRX* y un marcador microsatélite intragénico (D19S902) en las familias RP-1092 y RP-1192, con deleción de los exones 3 y 4 del gen. Análisis de segregación en la familia RP-0076, donde se observa penetrancia incompleta, siendo el paciente III:3 portador asintomático.

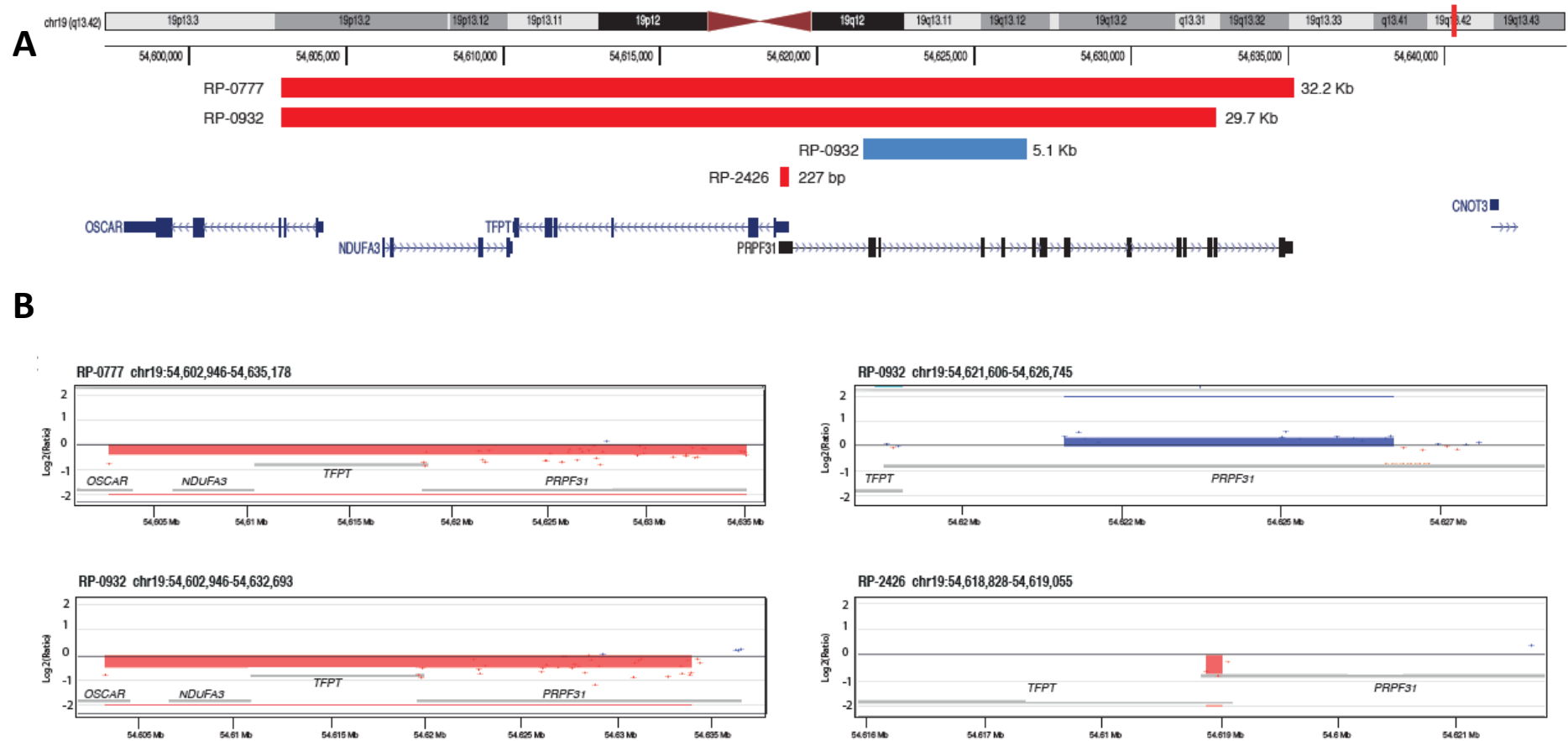


Figura 19. Acotamiento mediante aCGH de las deleciones del gen *PRPF31*, identificadas mediante MLPA o análisis de los datos de NGS. **A.** Representación de los reordenamientos en *PRPF31* y otros genes adyacentes en el cromosoma 19 en las familias RP-0777, RP-0932 y RP-2426 mediante aCGH. Las líneas rojas representan las posiciones genómicas de las deleciones en las familias RP-0777 (chr19: 54602946-54635178), RP-0932, paciente IV:12 (chr19: 54602946-54632693) y RP-2426 (chr19: 54618828-54619055). La línea azul indica las posiciones genómicas de la duplicación (chr19: 54621606-54626745) en el paciente IV:1 de la familia RP-0932. Se muestra una representación esquemática de la estructura completa del gen *PRPF31* y los genes contiguos. Los exones están indicados por rectángulos. **B.** Resultados del aCGH para las 4 familias. Se representa gráficamente la relación log2 del paciente frente a la referencia de cada sonda. Los puntos con razón de log2 alrededor de -1 indican una deleción heterocigota (puntos rojos), la razón de log2 alrededor de 0 indica un patrón normal, y +0,6 indica la amplificación heterocigótica (puntos azules). Las áreas sombreadas indican ganancias y pérdidas significativas. Las coordenadas genómicas se basan en hg19.

Los puntos de ruptura también se definieron para la familia RP-0777 usando el aCGH antes mencionado. Esta familia, que inicialmente fue negativa para el chip de genotipado de adRP, se caracterizó finalmente por WES (Almoguera *et al.*, 2015) con una gran deleción que abarca hasta el exón 13 del gen *PRPF31*, e incluye otros genes adyacentes (parte del gen *OSCAR* y los genes *NDUFA3* y *TFPT*). Por lo tanto, al realizar el aCGH identificamos una deleción mayor que la descrita, de al menos 32,2 kb (Tabla 8, Figura 19).

En las familias RP-1092 y RP-1192, con deleción de los exones 3 y 4 del gen *CRX* según el análisis de la profundidad de lectura de los datos de NGS, el aCGH de 60k en los *probandos* de ambas familias permitió determinar que presentaban deleciones con aparentemente distinto tamaño de 6,3 y 8 Kb, respectivamente (Tabla 8, Figura 20). Además, permitió ver que ningún otro gen estaba implicado en la región distal 3' de ambas deleciones.

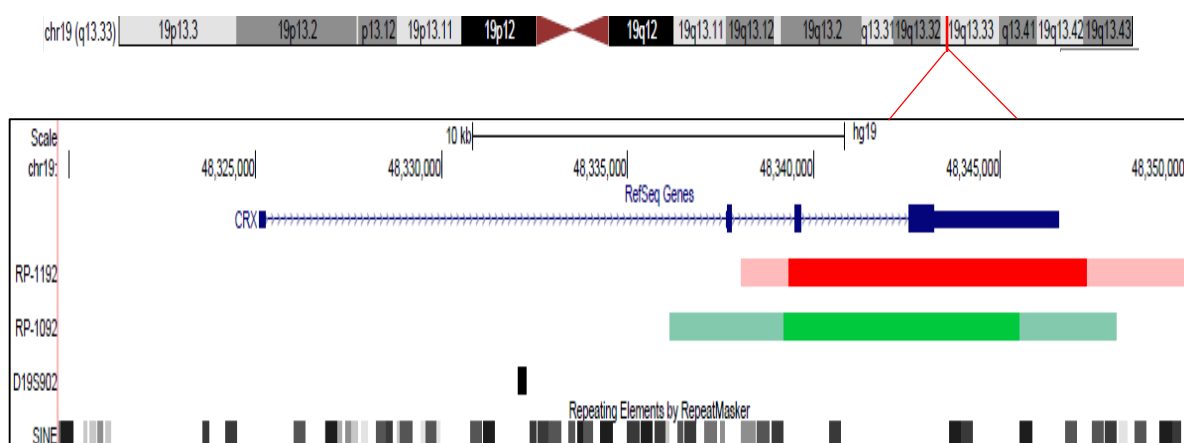


Figura 20. Acotamiento mediante aCGH de las deleciones del gen *CRX* en las familias RP-1192 y RP-1092, tras el análisis de reordenamiento usando los datos de NGS. La línea roja representa las posiciones genómicas de la deleción mínima de 8010 pb (chr19:48339345-48347355) y máxima de 11905 pb (chr19:48338052-48349957) en la familia RP-1192. La línea verde representa las posiciones genómicas de la deleción mínima y máxima de 6313 pb (chr19:48339216-48345529) y 12007 pb (chr19:48336134-48348141), respectivamente en la familia RP-1092. El marcador intragénico D19S902 está representado con un rectángulo negro. Se muestra la representación esquemática de la estructura completa de *CRX* en azul (los exones están representados con rectángulos) y de los elementos repetitivos (SINE) en negro.

4.3. Tasa global de detección de mutaciones en 258 familias con adRP en población española

Tras la aplicación secuencial de un cribado de SNVs y CNVs mediante distintos métodos en nuestra cohorte, 155 de las 258 familias obtuvieron un resultado genético concluyente. Todas ellas presentaron variantes claramente o probablemente patogénicas que fueron consideradas como la causa genética de la retinosis pigmentaria. Globalmente se obtuvo una tasa diagnóstica del 60,1% (155/258).

Con la finalidad de determinar cuál es el método más eficaz para el diagnóstico molecular de la adRP, en nuestra cohorte se realizó una comparativa de las tasas de caracterización genética entre las nuevas técnicas de secuenciación masiva con respecto a otras herramientas convencionales ("clásicas"), es decir, el uso combinado de SSCP/DGGE, microrray de genotipado y secuenciación Sanger. Tal como se detalla en la Tabla 9 y Figura 16, que resumen este análisis comparativo, la tasa diagnóstica claramente aumentó a lo largo del tiempo gracias a la implementación de la NGS. En nuestro laboratorio se ha pasado del 39% (87/226), usando el cribado molecular clásico, a una tasa de caracterización actual del 66% (21/32), obtenida en pacientes *naïve* que fueron estudiados mediante exoma clínico como primera aproximación diagnóstica. Esta nueva aproximación lleva dos años implementada en nuestro laboratorio (periodo 2015-2017), por lo que solo ha podido ser utilizada en un número limitado de familias reclutadas desde esta fecha.

Tabla 9. Descripción del rendimiento diagnóstico en función de la estrategia molecular seguida.

Algoritmo de estudio		Familias (No.)	Caracterizadas	No caracterizadas	Tasa Diagnóstica (%)
Clásico ^a	TOTAL	226	134	92	59,3
	Primer abordaje ^c	226	87	139	38,5
	NGS + CNVs Análisis secundario ^d	105	47	58	44,8
NGS ^b	Primer abordaje	32	21	11	65,6
TOTAL		258	155	103	60,1

Se muestra el número de familias caracterizadas respecto al número de familias estudiadas usando cada una de los abordajes moleculares. ^a Técnicas clásicas: SSCP/DGGE/Sanger y/o chip de genotipado. ^b Exoma clínico (*TruSightOne*) usando un panel virtual en familias *naïve* (no estudiadas previamente). ^c Familias *naïve* estudiadas usando técnicas clásicas. ^d Familias no *naïve* estudiadas con NGS (panel y/o WES) y/o análisis de CNVs.

Sin embargo, el número total de familias caracterizadas gracias a las técnicas de NGS en nuestra cohorte fue mayor y asciende a 60 casos, lo que representa el 23% (60/258) del total de familias y el 39% (60/155) de las familias caracterizadas. De ellos, 21 casos fueron caracterizados por exoma clínico, y otros 39 casos mediante captura con paneles personalizados, que fueron utilizados como segundo abordaje tras un cribado negativo mediante el abordaje clásico, anteriormente descrito. Debido a este sesgo, la tasa global de diagnóstico con las distintas herramientas de NGS respecto a las familias estudiadas con esta tecnología en nuestro laboratorio sería de 43,7% (60/137), que es inferior a la tasa obtenida al utilizar la NGS como primer abordaje debido a que 105 de las familias ya habían sido cribadas anteriormente con las técnicas tradicionales. Finalmente, gracias al análisis de CNVs se resolvieron 8 familias adicionales, en las que se identificaron deleciones o amplificaciones en genes asociados a adRP. En siguientes apartados se describe con más detalle el espectro mutacional de nuestra cohorte.

En un 40% (103/258) de las familias, los estudios genéticos realizados hasta la fecha no resultaron concluyentes, bien por la ausencia de detección de variantes patogénicas en los genes estudiados, o bien por la identificación de variantes de significado incierto. Cabe destacar que el 13% (34/258) de las familias de nuestra cohorte no pudieron ser estudiadas mediante NGS, debido a que no teníamos el consentimiento informado adecuado para poder realizar estudios de NGS, el *probandus* había fallecido, no querían seguir colaborando en el estudio, o el DNA era insuficiente o estaba degradado.

4.3.1. Espectro mutacional

Nuestra cohorte mostró una elevada heterogeneidad tanto alélica como genética. En total, se encontraron 160 variantes patogénicas que fueron consideradas como causales en las familias, de ellas 114 variantes fueron distintas: 106 variantes tipo SNVs y 8 CNVs (Tabla 8). El listado completo de las variantes SNVs se recoge en el Anexo 1.

Cincuenta y tres variantes han sido identificadas por primera vez en nuestra cohorte de pacientes españoles con adRP, lo que indica que las variantes asociadas a adRP suelen ser extremadamente raras. En este sentido, 89 variantes se encontraron en una única familia y por ello, el 78% (89/114) de las variantes identificadas representan eventos exclusivos de cada familia. En cambio, el 22% (25/114) de las variantes identificadas estaban presentes en dos o más familias, encontrándose en 71 familias (Anexo 1). Concretamente 18 variantes fueron identificadas en 2 familias y otras 7 variantes fueron consideradas recurrentes en nuestra cohorte por estar presentes en 3 o más familias no relacionadas, representando el 14% (36/258) y 13,6% (35/258) del total de las familias, respectivamente (Tabla 10). Entre ellas, la mutación

más frecuente en nuestra cohorte fue la variante p.Pro347Leu, localizada en el exón 5 del gen *RHO*, que representa un evento *hotstop* conocido en este gen. Esta mutación fue identificada en 10 familias no relacionadas (Fernandez-San Jose *et al.*, 2015b), lo que explica aproximadamente el 4% de la adRP en nuestra cohorte.

Tabla 10. Variantes patogénicas en 3 o más familias en nuestra población de 258 familias con adRP.

Gen	Exón	Cambio Nucleotídico	Cambio Proteína	No. Familias (%)
<i>RHO</i>	5	c.1040C>T	p.Pro347Leu	10 (3,9)
<i>NR2E3</i>	2	c.166G>A	p.Gly56Arg	7 (2,7)
<i>RHO</i>	2	c.403C>T	p.Arg135Trp	5 (1,9)
<i>C1QTNF5</i>	15	c.489C>A	p.Ser163Arg	4 (1,6)
<i>PRPF3</i>	11	c.1481C>T	p.Thr494Met	3 (1,2)
<i>RHO</i>	2	c.491C>A	p.Ala164Glu	3 (1,2)
<i>RHO</i>	3	c.544G>A	p.Gly182Ser	3 (1,2)
Total				35 (13,6)

Con respecto a la heterogeneidad genética observada en nuestra cohorte, se identificaron variantes causales en un total de 26 genes diferentes, asociados principalmente a formas dominantes de adRP, pero también a otras formas de DR (Tabla 11). Del total de los 24 genes asociados a adRP que fueron estudiados durante esta Tesis Doctoral, 17 genes presentaron alteraciones genéticas en 138 familias, lo que representa el 53,5% (138/258) del total de familias en estudio o el 89% (138/155) de las familias caracterizadas.

Los genes más frecuentemente mutados en nuestra cohorte fueron 4 genes asociados a adRP, en los que se identificaron variantes causales en el 62% (96/155) de las familias caracterizadas. De ellos, *RHO* fue el gen más prevalente, seguido de *PRPF31*, *RP1* y *PRPH2*, con prevalencias del 20,9%, 8,1%, 4,3% y 3,9%, respectivamente. En general, la mayoría de las variantes causales asociadas a adRP fueron de tipo *missense*, excepto para los genes *PRPF31*, *RP1*, *CRX* y *TOPORS* en los que las variantes de pérdida de función (LOF) fueron las más prevalentes (Tabla 11).

La frecuencia de CNVs en nuestra cohorte de adRP fue del 3,1% (8/258) y este tipo de evento mutacional representó el 5% (8/160) de todas las variantes identificadas. El gen *PRPF31* presentó la mayor frecuencia de CNVs en nuestra cohorte con 5 variantes diferentes identificadas en 4 familias (1,9%: 5/258), seguido de *CRX* (0,8%: 2/258) y *RHO* (0,4%: 1/258), (Tabla 8). Según el tipo de CNVs, la mayoría (75%, 6/8) fueron intragénicas, bien multiexónicas o afectando a un único exón. En las dos familias restantes, se identificaron deleciones de genes contiguos.

Tabla 11. Prevalencia de cada uno de los genes identificados en nuestra cohorte de 258 familias españolas con adRP, con el tipo de mutación en cada gen.

	Gen	No. familias	%	MS	FS	NS	SP	CNV	Otras
gen adRP	<i>RHO</i>	54	20,9	50	0	2	1	1	0
	<i>PRPF31</i>	21	8,1	2	5	2	6	5	1
	<i>RP1</i>	11	4,3	1	4	6	0	0	0
	<i>PRPH2</i>	10	3,9	8	0	0	2	0	0
	<i>NR2E3</i>	7	2,7	7	0	0	0	0	0
	<i>SNRNP200</i>	6	2,3	6	0	0	0	0	0
	<i>IMPDH1</i>	6	2,3	6	0	0	0	0	0
	<i>CRX</i>	5	1,9	1	2	0	0	2	0
	<i>NRL</i>	3	1,2	3	0	0	0	0	0
	<i>PRPF8</i>	3	1,2	1	0	0	0	0	2
	<i>PRPF3</i>	3	1,2	3	0	0	0	0	0
	<i>TOPORS</i>	3	1,2	0	1	2	0	0	0
	<i>BEST1</i>	2	0,8	1	0	0	0	0	0
	<i>GUCA1B</i>	1	0,4	1	0	0	0	0	0
	<i>HK1</i>	1	0,4	1	0	0	0	0	0
	<i>PROM1</i>	1	0,4	1	0	0	0	0	0
	<i>PRPF6</i>	1	0,4	0	0	0	1	0	0
gen adRD	<i>C1QTNF5</i>	4	1,6	4	0	0	0	0	0
Otro modelo de herencia	<i>RPGR</i>	4	1,6	0	4	0	0	0	0
	<i>CHM</i>	2	0,8	0	0	2	0	0	0
	<i>RP2</i>	2	0,8	0	0	0	0	0	2
	<i>ABCA4</i>	1	0,4	0	3	0	0	0	0
	<i>CNGA1</i>	1	0,4	0	1	1	0	0	0
	<i>EYS</i>	1	0,4	0	0	0	2	0	0
	<i>RDH5</i>	1	0,4	1	0	1	0	0	0
	<i>USH2A</i>	1	0,4	1	1	0	0	0	0
No caracterizada		103	39,9						
Total		258	100						

adRP: retinosis pigmentaria autosómica dominante; adRD: distrofia de retina autosómica dominante; MS: missense; FS: frameshift; SP: splicing.

Además, la utilización de la NGS nos permitió identificar una misma mutación en el gen *C1QTNF5*, asociado a una forma autosómica dominante de degeneración retiniana de inicio tardío, en 4 familias diferentes (1,6%: 4/258). Adicionalmente, se han identificado variantes causales en 3 genes asociados a RP ligada al cromosoma X, en 8 familias diferentes (3,1%: 8/258) y también en 5 genes asociados a varios tipos de DR con un patrón autosómico recesivo en 5 familias (1,9%: 5/258), gracias a la identificación de variantes bialélicas. Por lo que pudimos reclasificar clínica y genéticamente a un total de 6,6% (17/258) de las familias estudiadas.

4.3.2. Reclasificación clínica y genética de las familias con cambios recesivos o ligados al cromosoma X

El análisis de NGS reveló que 13 familias (5%) habían sido clasificadas erróneamente como autosómicas dominantes, el Anexo 2 muestra el árbol familiar de estas familias y los resultados del análisis de segregación. El análisis de esas familias mostró variantes causantes de enfermedad en genes ligados al cromosoma X, debido a la presencia de mujeres portadoras afectas: *RPGR* (4 familias), *RP2* (2 familias) y *CHM* (2 familias) (Tabla 12).

De forma similar, identificamos variantes patogénicas bialélicas compatibles con herencia autosómica recesiva en 5 familias, que incluyen una variante homocigótica de *splicing* en *EYS* y variantes heterocigóticas compuestas para los genes *RDH5*, *USH2A*, *CNGA1* y *ABCA4*. Las familias que portaban las variantes en los genes *RDH5* y *USH2A* solo se resolvieron parcialmente y, hasta el momento quedan por identificar variantes adicionales que expliquen la herencia pseudo-dominante en el resto de familiares afectados en la familia (Tabla 12, Anexo 2).

Se realizó una reevaluación clínica de los casos recesivos en los que los genes podrían asociarse con un diagnóstico clínico diferente al de RP. En la familia con variantes en *RDH5*, el *probandus* presentaba ceguera nocturna, ERG escotópico no registrable, papila y vasos de la retina normales, pequeños puntos blancos en la periferia media al polo posterior y ausencia de afectación macular en el examen del fondo de ojo. Estos hallazgos fueron compatibles, como el hallazgo genético, con una retinitis *punctata albescens*.

El *probandus* portador de variantes en el gen *USH2A* no presentaba pérdida auditiva y tampoco había casos de sordera en la historia familiar de esta manera, el fenotipo sugería RP en lugar de síndrome de Usher. En la familia con variantes en *ABCA4*, el fenotipo fue compatible con una distrofia de conos y bastones, con pérdida progresiva de agudeza visual seguida de constricción del campo visual y ceguera total a los 45 años de edad.

Tabla 12. Listado de SNVs probablemente patogénicas identificadas en nuestra cohorte en genes no asociados a adRP. En negrita: variantes descritas en nuestra población por primera vez.

Gen	Exón	Cambio Nucleotídico	Cambio Proteína	N Familias	Referencia
Gen con herencia autosómico dominante asociado a LORD (<i>late-onset retinal degeneration</i>)					
<i>C1QTNF5</i>	15	c.489C>A	p.Ser163Arg	4	(Hayward <i>et al.</i> , 2003)
Genes ligados al cromosoma X					
<i>CHM</i>	5	c.340G>T	p.Glu114*	1	Presente estudio
<i>CHM</i>	6	c.757C>T	p.Arg253*	1	(Hayakawa <i>et al.</i> , 1999)
<i>RP2</i>	1	c.1A>G	p.Met1?	1	Presente estudio
<i>RP2</i>	1	c.14_16del	p.Phe5del	1	(Neidhardt <i>et al.</i> , 2008)
<i>RPGR</i>	6	c.485_486del	p.Phe162Tyrfs*4	1	(Sharon <i>et al.</i> , 2000)
<i>RPGR</i>	8	c.888_889del	p.Ile297Lysfs*48	1	Presente estudio
<i>RPGR</i>	15	c.2405_2406del	p.Glu802Glyfs*32	1	(Vervoort <i>et al.</i> , 2000)
<i>RPGR</i>	15	c.2296_2299del	p.Gly766Asnfs*48	1	Presente estudio
Genes con herencia autosómica recesiva					
<i>ABCA4</i>	8	c.950del	p.Gly317Alafs*57	1	(Corton <i>et al.</i>, 2013)
	19	c.2888del	p.Gly963Alafs*14		(Paloma <i>et al.</i> , 2001)
	48	c.6688del	p.Leu2230Serfs*17		Presente estudio
<i>CNGA1</i>	5	c.94C>T	p.Arg32*	1	(Paloma <i>et al.</i> , 2002)
	6	c.131del	p.Glu44Glyfs*49		Presente estudio
<i>EYS</i>	IVS28	c.5928-2A>G	<i>Splicing</i>	1	(Gonzalez-del Pozo <i>et al.</i> , 2011)
<i>RDH5</i>	4	c.625C>T	p.Arg209*	1	(Schatz <i>et al.</i> , 2010)
	5	c.776C>T	p.Pro259Leu		Presente estudio
<i>USH2A</i>	13	c.2299del	p.Glu767Serfs*21	1	(Eudy <i>et al.</i> , 1998)
	62	c.12094G>A	p.Gly4032Arg		(Ge <i>et al.</i> , 2015)

4.4. Estudio del gen *PRPF31* en familias españolas con adRP

Las formas de adRP asociadas al gen *PRPF31* parecen tener unos rasgos distintivos con respecto a otros genes, principalmente por la presencia de un mayor número de reordenamientos y de familias con penetrancia incompleta, lo que dificulta la caracterización molecular y la toma de decisión sobre la posible causalidad de una nueva variante y por tanto, el consejo genético en las familias portadoras. Debido a que el gen *PRPF31* es el segundo gen más prevalente en nuestra cohorte de adRP afectando al 8,1% de las familias, en este trabajo nos propusimos profundizar en la implicación sobre la expresividad de nuevos mecanismos genéticos asociados a este gen, identificar posibles correlaciones genotipo-fenotipo, así como el estudio de posibles factores involucrados en la penetrancia incompleta asociada a la haploinsuficiencia en *PRPF31*.

4.4.1. Espectro mutacional de *PRPF31* en nuestra población española

Un total de 20 familias presentaron mutaciones puntuales o CNVs en el gen *PRPF31* tras un estudio genético mediante la combinación de las diferentes técnicas moleculares mencionadas anteriormente. En la Figura 21, se resume brevemente la secuencia del cribado de *PRPF31* en nuestra cohorte de 258 familias con adRP realizados a lo largo de las últimas dos décadas. Con este abordaje se caracterizaron:

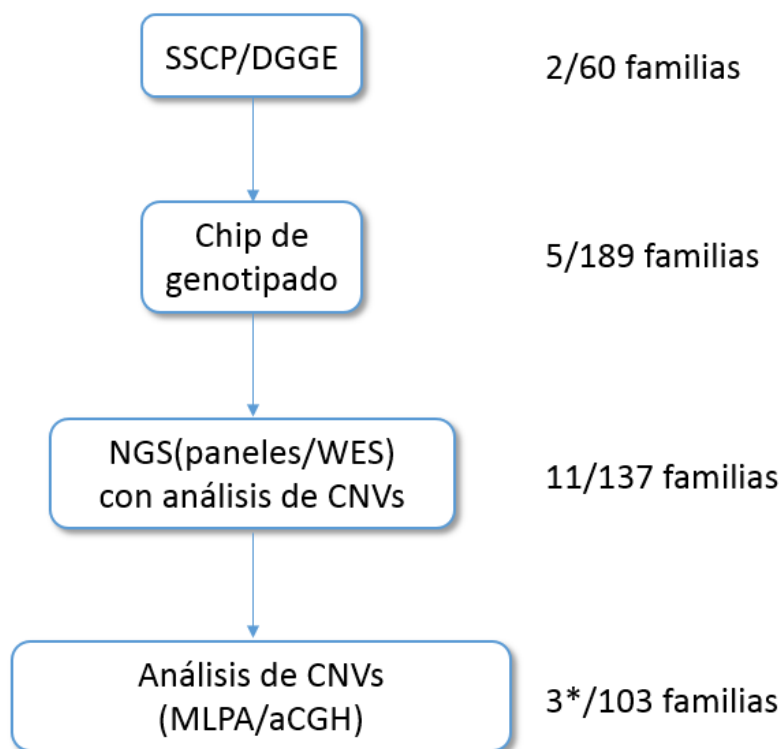


Figura 21. Estrategia molecular seguida para el cribado del gen *PRPF31* en una cohorte española de 258 familias con adRP. Al lado de cada técnica realizada se muestra el número de familias caracterizadas respecto al número de familias estudiadas. * Dos reordenamientos en una de las familias.

- i) 2 de las 60 familias estudiadas con SSCP y/o DGGE.
- ii) 5 de las 189 familias en las que se cribaron mutaciones conocidas en *PRPF31* mediante un chip de genotipado.
- iii) 11 familias gracias al uso de la NGS para estudio de SNVs y CNVs en 137 familias.
- iv) 2 familias adicionales portadoras de CNVs mediante MLPA.

Destacar que en una misma familia se identificaron dos reordenamientos genéticos diferentes segregando en distintas ramas familiares. Por lo tanto, se identificaron en total 21 mutaciones (Tabla 13), que incluían 6 mutaciones de *splicing* (28,6%), 5 mutaciones *frameshift* (23,8%), 4 deleciones de varios o todos los exones (19%), 2 mutaciones *nonsense* (9,5%), 2 mutaciones *missense* (9,5%), un cambio *in-frame* (4,8%) y una duplicación multiexónica (4,8%), siendo 8 de ellas nuevas. En total, los reordenamientos explican casi el 25% de los hallazgos en este gen. La prevalencia de mutaciones codificantes y reordenamientos genómicos en el gen *PRPF31* fue del 6,2% y 1,9%, respectivamente.

Las mutaciones en *PRPF31* parecen ser eventos exclusivos de cada familia, pues cada variante se identificó una única vez, con la excepción de dos mutaciones (p.Glu181* y c.322+4_322+7del), que estaban presentes en dos familias aparentemente no relacionadas. Un posterior análisis de haplotipos, utilizando marcadores microsatélites próximos al *locus* de *PRPF31* y/o SNPs intragénicos, concluyó que ambos casos se trataban de mutaciones con un origen ancestral común. En el caso de la mutación p.Glu181*, las familias RP-0006 y RP-1609 compartían una región genómica común flanqueando a la mutación, con un tamaño entre 1,2 - 1,4 Mb (Figura 22). En el caso de la mutación de *splicing* c.322+4_322+7del, el análisis de haplotipos en las familias RP-0187 y RP-1318 reveló la existencia de una región idéntica de al menos 1,7 Mb (Figura 23). Por lo tanto, ambas variantes parecen representar mutaciones fundadoras en nuestra población en lugar de tratarse de eventos recurrentes.

El análisis de segregación de las variantes causales en *PRPF31* en sus respectivas familias, recogido en el Anexo 3, nos permitió identificar que un 65% (13/20) de las familias mostraban penetrancia incompleta (Tabla 13). Por lo tanto, en 13 familias existían uno o varios portadores asintomáticos, según los datos oftalmológicos e historia familiar disponibles en el momento del estudio.

Tabla 13. Mutaciones en el gen *PRPF31* encontradas en la cohorte de 258 familias españolas con adRP. ND: no hay datos disponibles.

Familia	Mutaciones en el gen <i>PRPF31</i>						
	Exón	Cambio Nucleotídico	Cambio Proteína	Referencia	Técnica	Penetrancia Incompleta	Segregación
RP-0932_del	Delección de <i>OSCAR</i> (E1 al E2), <i>NDUFA3</i> , <i>TFPT</i> y <i>PRPF31</i> (E1 al 13)			Presente estudio	MLPA	Si	Si
RP-0932_dup	Duplicación de <i>PRPF31</i> (E2 al E5)			Presente estudio	MLPA	Si	Si
RP-2426	Delección de <i>PRPF31</i> (E1)			Presente estudio	MLPA	ND	ND
RP-0076	Delección de <i>PRPF31</i> (E9)			Presente estudio	panel NGS	Si	Si
RP-2712	2	c.55del	p.Glu19Lysfs*46	Presente estudio	panel NGS	ND	ND
RP-2496	IVS10-1	c.1074-1G>T	<i>Splicing</i>	Presente estudio	panel NGS	No	ND
RP-0182	IVS11+2	c.1146+2T>A	<i>Splicing</i>	Presente estudio	adRP Chip	Si	Si
RP-2266	14	c.1462_1472del	p.Lys488Argfs*75	Presente estudio	panel NGS	Si	ND
RP-0777	Delección de <i>OSCAR</i> (E1 al E2), <i>NDUFA3</i> , <i>TFPT</i> y <i>PRPF31</i> (E1 al 14)			(Almoguera <i>et al.</i> , 2015)	WES	No	Si
RP-0187	4	c.322+4_322+7del	<i>Splicing</i>	(Q. Zhang <i>et al.</i> , 2016)	panel NGS	Si	Si
RP-1318	4	c.322+4_322+7del	<i>Splicing</i>	(Q. Zhang <i>et al.</i> , 2016)	panel NGS	Si	ND
RP-0361	5	c.328_330del	p.Ile110del	(de Sousa Dias <i>et al.</i> , 2013)	panel NGS	Si	Si
RP-2136	IVS6+1	c.527+1G>T	<i>Splicing</i>	(Chakarova <i>et al.</i> , 2006)	adRP Chip	No	ND
RP-0006	7	c.541G>T	p.Glu181*	(Pomares <i>et al.</i> , 2010)	adRP Chip	No	Si
RP-1609	7	c.541G>T	p.Glu181*	(Pomares <i>et al.</i> , 2010)	adRP Chip	Si	Si
RP-2595	8	c.736G>A	p.Ala246Thr	(Xu <i>et al.</i> , 2014)	panel NGS	ND	ND
RP-0637	8	c.770dup	p.Thr258Aspfs*21	(Vithana <i>et al.</i> , 2001)	SSCP/DGGE	Si	Si
RP-0368	8	c.828_829del	p.His276Glnfs*2	(Martinez-Gimeno <i>et al.</i> , 2003)	SSCP/DGGE	Si	Si
RP-1599	9	c.895T>C	p.Cys299Arg	(Sullivan <i>et al.</i> , 2006a)	adRP Chip	Si	Si
RP-1541	9	c.939dup	p.Gly314Argfs*10	(Fernandez-San Jose <i>et al.</i> , 2015a)	panel NGS	Si	Si
RP-1970	IVS11+2	c.1146+2T>C	<i>Splicing</i>	(Waseem <i>et al.</i> , 2007)	panel NGS	Si	Si

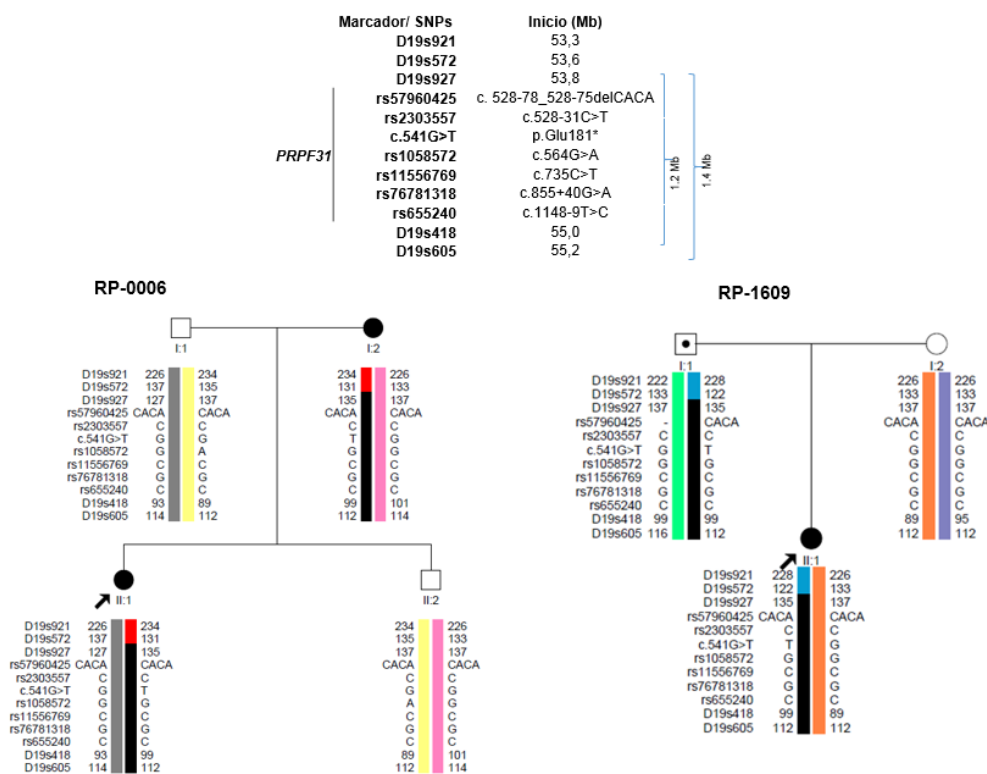


Figura 22. Análisis de haplotipos en dos familias con la misma variante nonsense en *PRPF31*. Para ello se utilizaron 6 SNPs intragénicos y 5 marcadores microsatélites que flanquean *PRPF31* en las familias RP-0006 y RP-1609 con la misma mutación causal p.Glu181*. El análisis de los haplotipos mostró una región común que compartía un mínimo de 1,2 Mb y un máximo de 1,4 Mb definido por los alelos ubicados a lo largo de la barra negra (el marcador D19S605 no fue informativo).

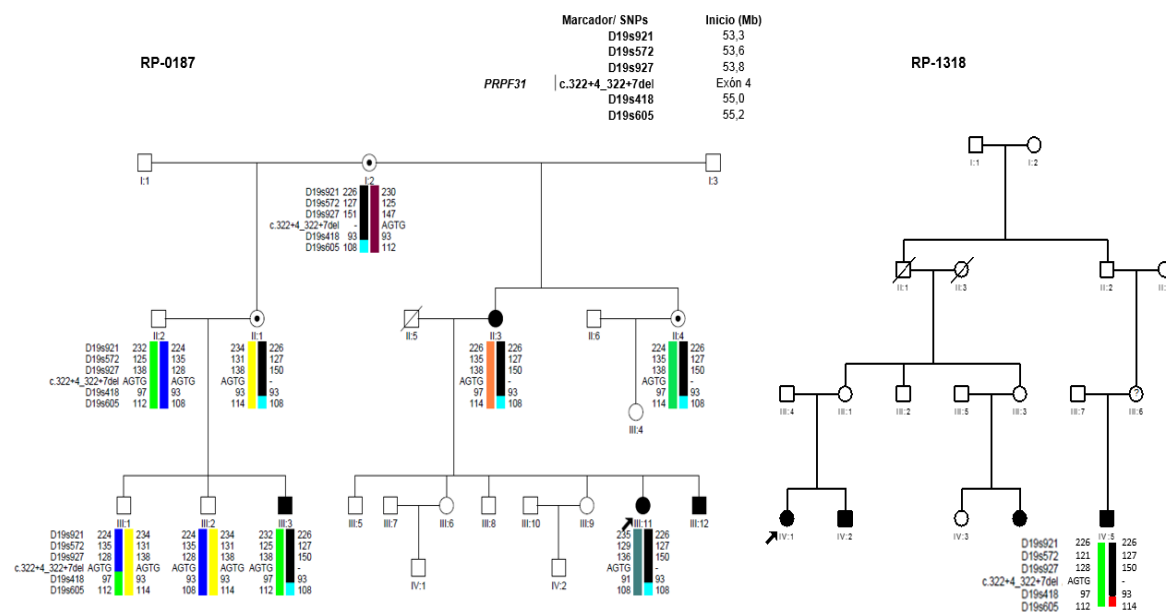


Figura 23. Análisis de haplotipos en dos familias con la misma variante de *splicing* en *PRPF31*. Se usaron 5 marcadores microsatélites que flanquean *PRPF31* en las familias RP-0187 y RP-1318 con la misma mutación causal c.322+4_322+7del. El análisis de los haplotipos mostró una región común que compartía un mínimo de 1,7 Mb definido por los alelos ubicados a lo largo de la barra negra.

4.4.2. Dos reordenamientos en *PRPF31* segregando en la misma familia

En la familia RP-0932 que presentaba un gran árbol familiar en donde la RP se encontraba segregando en 9 individuos de 2 ramas familiares distintas (Figura 17), observamos dos reordenamientos genéticos diferentes en el gen *PRPF31*. En concreto, se identificó una duplicación de los exones 2 a 5 en el caso índice y en su padre afectado (IV:1 y III:1, respectivamente), y una gran deleción de los exones 1 a 13 en otros dos individuos afectados de la familia, su tía paterna y su prima (III:12 y IV:12, respectivamente).

Este caso presenta dos hallazgos novedosos en la etiopatogenia de la RP pues es la primera vez que se identifica una duplicación en el gen *PRPF31* como probable causa de patología y además, la presentación de dos reordenamientos diferentes dentro de la misma familia, es un fenómeno no descrito anteriormente en ninguna otra familia con RP. Con el fin de estudiar la implicación de estos reordenamientos en una desregulación de *PRPF31*, se realizó un análisis de expresión génica mediante qPCR en distintos portadores de ambas CNVs. Tal como se describe en la Figura 24, los niveles de expresión de *PRPF31* disminuyeron significativamente en los pacientes con la deleción (IV:12) y/o la duplicación (IV:1 y III:1) con un ratio entre 0,52 y 0,7 en comparación con muestras de controles ($p < 0,001$). Por lo tanto, estos pacientes mostraron una disminución en los niveles de expresión génica de hasta el 48% del valor normal, lo que apoya que la haploinsuficiencia de *PRPF31* sea el mecanismo patogénico responsable de la RP, asociado no solo a la deleción sino también a la duplicación parcial del gen.

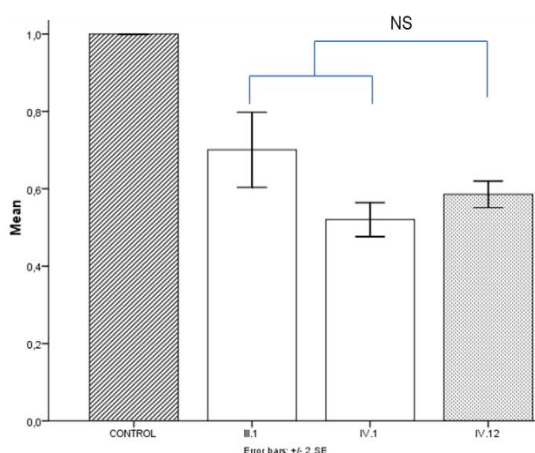


Figura 24. Expresión del gen *PRPF31* en la familia RP-0932. La expresión de *PRPF31* disminuyó en los pacientes con la duplicación (III:1 y IV:1) y la deleción (IV:12), en comparación con los controles, siendo las diferencias de las medias entre ambos grupos estadísticamente significativas (valor de $p < 0,001$), mientras que las diferencias de las medias entre los pacientes con la duplicación y el paciente con la deleción no fueron estadísticamente significativas (NS). El ensayo se realizó por triplicados dos veces.

Esta amplificación, tras su detección mediante MLPA, fue posteriormente verificada y acotada por aCGH, tal como describimos en el apartado 4.2.3. Con la finalidad de determinar la localización y reorganización, se utilizó una estrategia de *long-Range* PCR utilizando cebadores específicos. Como se observa en la Figura 25, los pacientes IV:1 y III:1 portaban la región génica amplificada a continuación de los exones 2-5, lo que confirma una duplicación en *tándem*. Este

hallazgo es consistente con la pérdida de expresión del alelo mutante, probablemente debido a que la duplicación de estos 4 exones conduzca a la producción de isoformas aberrantes que sean degradadas mediante el sistema *non-sense mediated decay* (NMD).

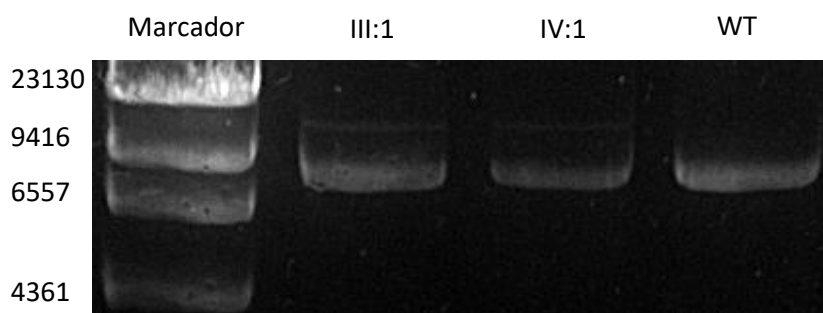


Figura 25. Validación mediante PCR larga de la duplicación en tándem en la familia RP-0932. Gel de agarosa al 1%. Como marcador se usó Lambda DNA / HindIII. Tamaño del marcador en pb. Se observó una banda de 8,6 kb, correspondiente al alelo normal, en los pacientes III:1, IV:1 y la muestra control (WT). Los pacientes III:1 y IV:1 además presentaban una banda superior de entre 9,4 a 23,1 kb, correspondiente al alelo duplicado.

4.4.3. Contribución del elemento repetido MSR1 a la penetrancia de la enfermedad de las mutaciones en *PRPF31*

Debido a la alta frecuencia de penetrancia incompleta asociada a *PRPF31* en nuestra cohorte, en este trabajo nos planteamos profundizar en la identificación de otros factores genéticos en nuestras familias que pudieran modular las diferencias en la expresión del alelo mutante en pacientes y portadores asintomáticos.

Recientemente, Rose y colaboradores (Rose *et al.*, 2016) describieron que variaciones en el número de repeticiones del minisatélite MSR1 situado aguas arriba del promotor del gen *PRPF31*, podrían tener un papel fundamental en la penetrancia de las mutaciones en *PRPF31*, al actuar como modulador de su actividad transcripcional. Por ello, decidimos estudiar la implicación de este elemento repetitivo MSR1 en la expresividad de la enfermedad en nuestra cohorte. Se analizó el número de copias de este MSR1 en 40 portadores de mutaciones en *PRPF31*: 35 pacientes sintomáticos de 15 familias diferentes y 5 individuos asintomáticos de 5 de estas familias (Anexo 3).

Tras el genotipado, el 54% de los individuos sintomáticos (19/35) portaba el genotipo 3/4 de MSR1 y solo el 46% (16/35) portaba el genotipo 3/3. En este último grupo, se encuentran cuatro individuos hemicigotos para el alelo 3 de MSR1 debido a la presencia de grandes deleciones de *PRPF31* que afectan a la región de MSR1. Ningún paciente afectado presentaba el genotipo 4/4. Esta alta proporción del alelo de 4 copias de MSR1 en pacientes sintomáticos varía con respecto a la frecuencia descrita anteriormente en el trabajo de Rose (Rose *et al.*, 2016), en donde todos

los pacientes sintomáticos fueron 3/3. Con el fin de determinar si estas diferencias podían ser poblacionales, se estimó la MAF en población española, siendo la frecuencia de este alelo 0,13, similar a la reportada anteriormente de 0,154 (Rose *et al.*, 2016). En el grupo de individuos asintomáticos se identificaron los genotipos 3/3 (20%), 3/4 (40%) y 4/4 (40%), estando el alelo de 4 copias de MSR1 significativamente sobrerrepresentado (valor $p < 0,001$) con respecto a los controles.

Estas observaciones están en concordancia con la hipótesis de que el alelo de 4 copias proporciona una mayor expresión del gen *PRPF31* y podría actuar como un factor protector en la expresividad de la enfermedad asociada a *PRPF31*. La penetrancia incompleta se observó en todas las familias en donde los pacientes sintomáticos presentaban el genotipo 3/4 para MSR1 (Anexo 3).

4.4.4. Correlación genotipo-fenotipo para las familias con variantes en *PRPF31*

En la Tabla 14 y Tabla 15 se presentan un resumen de los datos oftalmológicos de 31 pacientes de 18 familias afectadas de RP por mutaciones en *PRPF31*, así como un análisis de las características clínicas en función del tipo de mutación.

En general, los pacientes presentaron ceguera nocturna como primer síntoma (edad media $13,4 \pm 7,6$ años), seguido de constricción del campo visual ($18,5 \pm 9,9$ años) y pérdida de la agudeza visual ($25,2 \pm 12,4$ años). En general, en aquellos pacientes en los que se disponía de datos de ERG, la actividad de los fotorreceptores no pudo ser registrada. Hasta el 36% de los pacientes presentaron cataratas subcapsulares, con un rango de aparición entre los 21 a los 65 años. En el examen funduscópico, los pacientes presentaron cambios típicos asociados a la RP, principalmente dispersión de pigmento, estrechamiento de vasos y palidez papilar. En algunos casos, se presentaron alteraciones maculares en pacientes con una pérdida importante de la agudeza visual.

Aparentemente, los síntomas más severos se observaron en el paciente portador de la duplicación en tándem (individuo IV:1, familia RP-0932, Figura 17), en términos de inicio y evolución de la enfermedad. Por otro lado, los casos con mutaciones tipo *missense/in-frame* parecen tener un inicio más temprano y una evolución rápida a la pérdida de la visión. En las familias en las que la delección de *PRPF31* incluía la delección de genes flanqueantes (genes *OSCAR*, *NDUFA3* y *TFPT*) no se asoció con signos adicionales aparte de la RP.

Tabla 14. Características clínicas de 31 pacientes afectados de 18 familias caracterizadas con una mutación en el gen *PRPF31*. Fondo de RP típico: palidez papilar, atenuación de los vasos de la retina y cambios pigmentarios (espículas). AM: alteración macular. CN: ceguera nocturna. CCV: constricción del campo visual. AV: agudeza visual. BCVA: agudeza visual corregida. CF: contar dedos. OD: ojo derecho. OI: ojo izquierdo. ERG: electroretinograma. NR: no registrable. ND: no hay datos disponibles. Esco: escotópico. Fot: fotópico. RA: amplitud reducida. a.:años.

Familia	Mutaciones <i>PRPF31</i>	ID	Primeros síntomas (años)			Edad examinación	BCVA (OD/OI)	Campo visual	ERG	Fondo de ojo	Otros hallazgos
			CN	CCV	AV						
RP-0006	c.541 G>T	VI:5	8	9	No	44	0,2/0,3	Tubular	NR	RP típico	Cataratas (42a.)
		VII:9	5	Si	Si	19	0,5/0,5	Tubular	NR	RP típico	
RP-0076	del chr19:54629903-54629992	III:2	30	36	36	54	0,1/0,1	Tubular	NR	RP típico	Fotofobia. Cataratas
		IV:2	15	No	No	21	0,5/0,5	Escotoma anular	Esco:NR.	RP típico	
RP-0182	c.1146+2T>A	IV:4	7	Si	Si	40	0,3/0,3	Tubular	NA	RP típico	Cataratas (38a.)
RP-0187	c.322+4_322+7del	III:11	25	25	26	30	0,4/0,5	Escotoma anular	Esco:NR. Fot: RA	RP típico	
		II:3	no	Si	No	52	0,7/0,8	Tubular	NA	RP típico	
RP-0361	c.328_330del	IV:1	14	14	20	27	0,9/0,9	Tubular	NR	RP típico	
		IV:6	14	14	20	27	0,7/0,7	Tubular	NA	RP típico	
		IV:7	Si	Si	Si	24	0,7/0,6	Tubular	NA	RP típico	
RP-0368	c.828_829del	II:1	20	20	Si	56	NA	NA	NA	NA	Cataratas (56a.)
		III:3	Si	Si	No	39	1/0,6	Tubular	NA	RP típico	
RP-0637	c.770dup	II:2	18	40	50	64	CD/CD	Tubular	NR	RP típico + AM	Fotofobia. Discromatopsia. OD Cataratas (65a.)
		III:1	10	20	35	40	0,7/0,7	Tubular	NR	RP típico	OI Cataratas (38a.)
		IV:1	NA	NA	NA	16	0,7/1	Constricción periférica + pérdida difusa	NA	Edema macular cistoide	

Tabla 14. Características clínicas de 31 pacientes afectados de 18 familias caracterizadas con una mutación en el gen *PRPF31* (Continuación).

Familia	Mutaciones <i>PRPF31</i>	ID	Primeros síntomas (años)			Edad examinación	BCVA (OD/OI)	Campo visual	ERG	Fondo de ojo	Otros hallazgos
			CN	CCV	AV						
RP-0777	del chr19:54602946-54635178	III:1	27	32	No	37	0,7/0,8	Tubular	NR	RP típico	
RP-0932	dup chr19:54621606-54626745	III:1	Si	Si	Si	A edad adulta	NA	NA	NA	NA	
		IV:1	5	7	20	30	0,2/0,2	Tubular	NA	RP típico + AM	Discromatopsia. Cataratas (25a.)
	del chr19:54602946-54632693	III:12	Si	14	13	51	0,125/0,125	Tubular	NA	RP típico + AM	Cataratas (50a.)
		IV:12	Si	17	No	28	0,8/0,8	Tubular	NA	RP típico	
RP-1318	c.322+4_322+7del	III:1	9	9	16	NA	NA	NA	NA	NA	
RP-1541	c.939dup	IV:2	6	20	39	38	0,1/0,1	NA	NA	RP típico	Cataratas (21a.)
RP-1599	c.895T>C	III:6	Si	7	7	43	0,1/0,05	Tubular	NA	RP típico	Cataratas (30a.)
		IV:4	3	5	6	15	0,8/0,6	NA	NA	RP típico	
RP-1609	c.541G>T	IV:2	6	22	25	23	0,8/0,7	Tubular	Esco:NR. Fot: RA	Cambios pigmentarios periféricos	Fotofobia. Discromatopsia
RP-1970	c.1146+2T>C	V:1	14	12	No	16	NA	Constricción periférica 30º	Esco:NR. Fot: RA	NA	Hipermetropía. Astigmatismo
		V:2	13	No	No	14	1,0/1,0	Constricción periférica 30º	Esco:NR. Fot: RA	Cambios pigmentarios + estrechamiento vasos	
RP-2136	c.527+1G>T	II:1	Si	Si	Si	41	0,2/0,2	Tubular	NR	RP típico	

Tabla 14. Características clínicas de 31 pacientes afectados de 18 familias caracterizadas con una mutación en el gen *PRPF31* (Continuación).

Familia	Mutaciones <i>PRPF31</i>	ID	Primeros síntomas (años)			Edad examinación	BCVA (OD/OI)	Campo visual	ERG	Fondo de ojo	Otros hallazgos
			CN	CCV	AV						
RP-2266	c.1462_1472del	III:1	13	16	30	51	0,3/0,3	Tubular	NA	RP típico	Cataratas (50a.)
RP-2426	del chr19:54618828-54619055	III:1	NA	NA	NA	37	0,8/0,8	Tubular	NA	RP típico	
RP-2496	c.1074-1G>T	III:2	20	30	35	36	0,6/0,8	Tubular	NA	RP típico	Fotofobia

Tabla 15. Descripción del fenotipo en función de la edad de aparición de los síntomas, en pacientes con mutación en *PRPF31*. Edad de inicio de la ceguera nocturna (CN), constricción del campo visual (CV), pérdida de la agudeza visual (AV), cataratas y porcentaje de individuos con alteración de la AV y cataratas según el tipo de mutación. Los datos se muestran como la media \pm desviación estándar. FS: *frameshift*. No: número. Entre paréntesis se muestra el número de individuos con datos de la edad de inicio de los síntomas en cada una de las distintas mutaciones.

Tipo de mutación	CN (No. casos)	CV (No. casos)	AV (No. casos)	% AV	Cataratas (No. casos)	% Cataratas
<i>Splicing</i>	14,7 \pm 6,8 (6)	19,0 \pm 10,1 (4)	25,7 \pm 9,5 (3)	62,5%	38,0 (1)	13%
<i>Missense/in-frame</i>	10,3 \pm 6,4 (3)	10,0 \pm 4,7 (4)	13,3 \pm 7,8 (4)	100%	30,0 (1)	20%
<i>FS/nonsense</i>	10,8 \pm 5,7 (8)	21,0 \pm 9,4 (7)	35,8 \pm 9,5 (5)	77,8%	45,3 \pm 15,4 (6)	60%
Duplicación	5,0 (1)	7,0 (1)	20,0 (1)	100%	25,0 (1)	50%
Delección	24 \pm 7,9 (3)	24,8 \pm 10,9 (4)	24,5 \pm 16,3 (2)	40%	50,0 (1)	33%
TOTAL	13,4 \pm 7,6	18,5 \pm 9,9	25,2 \pm 12,4		37,7 \pm 10,4	

Discusión

El objetivo principal de esta Tesis ha sido profundizar en el conocimiento sobre las bases genéticas y moleculares de pacientes con retinosis pigmentaria autosómica dominante. Para ello, se han evaluado y utilizado de forma secuencial diferentes herramientas moleculares, tanto clásicas como varios abordajes de secuenciación masiva, así como el estudio de la implicación de CNVs en la etiopatogenia de esta forma de RP. Dado el exhaustivo estudio realizado en estas familias, este trabajo aporta prevalencias actualizadas de mutaciones y de los genes responsables de la enfermedad, lo que permite conocer la implicación real de al menos 24 de los 28 genes asociados a adRP. Además, hemos implementado un algoritmo de diagnóstico genético coste efectivo para el estudio de esta patología. Por último, se ha profundizado en las correlaciones genotipo-fenotipo del segundo gen más prevalente en nuestra cohorte, identificado nuevos reordenamientos, además de aportar luz en los mecanismos que pueden estar asociados a la penetrancia incompleta asociada a *PRPF31*.

5.1. Datos globales de la caracterización genética de la adRP en España

En este trabajo se presenta el estudio de la mayor cohorte de familias en población europea estudiada hasta la fecha, con un diagnóstico inicial de RP con sospecha de herencia autosómica dominante, lo que ha permitido establecer la tasa de caracterización actual de adRP en población española, así como la prevalencia de los distintos genes en esta población y los eventos mutacionales encontrados en la misma. Esta cohorte de 285 familias ha sido estudiada a lo largo de 26 años aplicando diferentes estrategias disponibles en cada época en nuestro laboratorio, lo que aporta una visión muy interesante de cómo la tasa diagnóstica en adRP ha ido cambiando a lo largo del tiempo, gracias al desarrollo de nuevas estrategias de diagnóstico genético. Actualmente la secuenciación masiva parece ser la aproximación más rentable para el diagnóstico molecular de pacientes con adRP (Bowne *et al.*, 2011; Bernardis *et al.*, 2016), tal y como también hemos demostrado en nuestro estudio.

Usando las técnicas de cribado SSCP/DGGE/Sanger y/o chip de genotipado de los genes más frecuentemente mutados se obtuvo una tasa diagnóstica global del 38%. Tras realizar análisis de NGS y estudios de CNVs, la tasa diagnóstica total aumentó al 60,1%, siendo superior a la encontrada en población alemana (41%) (Glockle *et al.*, 2014) o belga (56%) (Van Cauwenbergh *et al.*, 2017), pero menor que la tasa del 78,5% descrita en una cohorte de Norte América (Daiger *et al.*, 2014a), la mayor cohorte estudiada hasta la fecha. La presencia de la mutación fundadora p.Pro23His en el gen *RHO*, exclusiva de los Estados Unidos, en el 13,2% de las familias adRP americanas (Daiger *et al.*, 2014a), puede explicar la diferencia observadas entre ambas poblaciones. Al no haberse identificado ninguna mutación con tan alta prevalencia, nuestra

cohorte presenta una mayor diversidad alélica que la observada en EE.UU. Además, la tasa diagnóstica obtenida en nuestra cohorte está de alguna manera sesgada e infraestimada, ya que 34 de las 258 familias adRP (13,2%) no se pudieron estudiar usando NGS debido a la falta de disponibilidad de nuevas muestras de ADN con la calidad suficiente en familias antiguas y/o fallecimiento de los pacientes en estudio, fundamentalmente.

En este estudio se realizó una comparación de la capacidad diagnóstica de la NGS en dos subgrupos de casos: los casos retrospectivos no caracterizados con técnicas clásicas, y un subgrupo de nuevas familias recopiladas de forma prospectiva durante el desarrollo de esta Tesis Doctoral, en los que la NGS fue utilizada como primera aproximación diagnóstica, utilizando la estrategia de exoma clínico con prefiltrado de genes asociados a DR. Así, teniendo en cuenta que la tasa diagnóstica en estos casos *naïve* fue de un 66%, es probable que la tasa diagnóstica actual en familias con adRP se encuentre entre el 60-70%.

Durante este trabajo se han evaluado otras estrategias de NGS a lo largo de los últimos 5 años. Se han utilizado sistemas mixtos de captura de regiones candidatas mediante paneles customizados, diseñados en exclusiva para el estudio de DR no sindrómicas, mediante una aproximación comercial usando HaloPlex (con 75 genes), como una herramienta no comercial basada en *Molecular Inversion Probes* (con 108 genes) desarrollada en colaboración con el *Raboud University Medical Center*. En total, se caracterizaron 43 pacientes con una tasa diagnóstica entre el 40-51%, cifras inferiores a las obtenidas en los casos *naïve*. En otros trabajos sobre el uso de la NGS en pacientes con adRP también muestran una gran variabilidad en el rendimiento de detección de mutaciones, entre el 23 y 86% (Bowne *et al.*, 2011; Audo *et al.*, 2012; Eisenberger *et al.*, 2013; Daiger *et al.*, 2014b; Glockle *et al.*, 2014), similar a lo que observamos en nuestro estudio utilizando distintas estrategias. Esta variabilidad en las tasas de detección se debe a diferencias en las características de las subcohortes estudiadas, en el número de genes incluidos en los diferentes paneles y principalmente, en los estudios moleculares previos realizados en los pacientes. En nuestro caso, 105 de las 137 familias analizadas por NGS fueron analizadas previamente mediante el *array* de genotipado de adRP, cribado de *RHO* (Fernandez-San Jose *et al.*, 2015b), *PRPH2* (Trujillo *et al.*, 1998), y de exones concretos de los genes *NR2E3* (Blanco-Kelly *et al.*, 2016), *PRPF3*, *PRPF8* (Martinez-Gimeno *et al.*, 2003), *RP1* (Gamundi *et al.*, 2006) y *SNRNP200*, lo que probablemente subestimó la tasa de detección en casos no *naïve*.

Recientemente, se ha evidenciado que las CNVs son una causa importante e infraestimada de DR (Ellingford *et al.*, 2017; Van Schil *et al.*, 2017). Se han descrito delecciones que implican varios

exones en distintos genes asociados a DR en un importante número de pacientes, no caracterizados previamente tras análisis de NGS (Ellingford *et al.*, 2017). Nuestro estudio es pionero en la implementación del análisis de CNVs para la caracterización molecular en pacientes con adRP, lo que nos ha permitido subrayar la importancia que tienen las variantes estructurales en la etiología de la adRP. Primero se evaluaron deleciones en los genes más prevalentes, principalmente *RHO* y *PRPF31*, utilizando un kit comercial de MLPA. Además, quisimos evaluar la capacidad de detección de CNVs aprovechando los datos de NGS ya disponibles en 68 casos que fueron obtenidos con la estrategia de HaloPlex. Su elección frente a otras disponibles en el laboratorio se debió a dos razones principales como la disponibilidad de un mayor número de casos analizados con esta técnica, además de su análisis a alta profundidad (>300x) y con muy buena cobertura de las regiones capturadas (99,1%). Esto nos aseguró la obtención de datos más homogéneos y robustos durante la comparativa de las profundidades de secuencia, con el fin de minimizar falsos positivos debidos a variabilidad de cobertura en las regiones capturadas y/o errores de secuenciación.

Para ello, se reanalizaron los datos crudos (*raw data*) de secuencia utilizando una nueva *pipeline* para analizar SNVs y CNVs. En el caso de las SNVs, se utilizaron algoritmos más actualizados, nuevas herramientas de anotación y predicción de patogenicidad, así como nuevas bases de datos de frecuencias. Como consecuencia de este análisis, no solo se identificaron deleciones afectando uno o varios exones en 3 pacientes, sino que además se encontraron 5 nuevas SNVs causales que no habían sido detectadas con el análisis bioinformático anterior. Esto permitió que el rendimiento de esta estrategia aumentara del 27% descrito anteriormente (Fernandez-San Jose *et al.*, 2015a) al 39,7%. Nuestro trabajo pone de manifiesto la utilidad de realizar reanálisis sistemático de los datos de NGS a medida que mejora la sensibilidad de los algoritmos bioinformáticos y aumenta la información de las bases de datos.

Cabe destacar que el 3,1% de nuestros pacientes son portadores de CNVs de distinto tamaño, principalmente en los genes *PRPF31* y *CRX*. Estas CNVs fueron identificadas mediante MLPA y/o análisis de NGS, confirmadas mediante un análisis de aCGH para definir mejor el tamaño, acotar los puntos de corte e identificar otros posibles genes adyacentes que pudiesen estar involucrados. De hecho, el uso del aCGH nos permitió refinar los puntos de rotura de los reordenamientos de *PRPF31* y *CRX*, aunque debido a una baja cobertura de las regiones adyacentes a estos genes en los aCGH utilizados no pudimos establecer los puntos de rotura exactos para las deleciones de mayor tamaño identificadas en nuestros pacientes. Este fue el caso de las deleciones en *PRPF31* encontradas en las familias RP-0932 y RP-0777, que incluyen genes contiguos, y donde la deleción parece empezar en el mismo punto, extremo 5', en ambas

familias. Por otro lado, en la familia RP-0777 portadora de una delección identificada mediante WES, la confirmación mediante aCGH evidenció que tenía un mayor tamaño que el descrito previamente (Almoguera *et al.*, 2015), incluyendo el exón 14 y el exón 2 de los genes *PRPF31* y *OSCAR*, respectivamente.

Esta elevada frecuencia de CNVs para dichos genes puede deberse a su localización genómica en una zona rica en elementos repetidos, especialmente de repeticiones tipo Alu. Los elementos Alu son un tipo de SINE (*Short Interspersed Element*) que constituye el 10% del contenido total del genoma, pero que constituye el 26,3% del cromosoma 19, siendo el cromosoma más rico en secuencias tipo Alu (Grover *et al.*, 2004) (Figura 26). Las repeticiones tipo Alu están implicadas en enfermedades genéticas humanas por inserción de estos elementos en el genoma o por recombinación homóloga no alélica causando reordenamientos. Este último mecanismo, se predice que causa el 0,3% de las enfermedades genéticas humanas (Deininger & Batzer, 1999), y puede ser la explicación más probable de la alta prevalencia de reordenamientos genómicos en el gen *PRPF31* (Rose *et al.*, 2011). Por ello, el mecanismo subyacente a los complejos reordenamientos estructurales en estas familias, con delecciones tanto en *PRPF31* como en *CRX*, podría ser por recombinación homóloga no alélica entre dos repeticiones tipo Alu con elevada homología, presentes en los extremos 5' y 3' de las regiones delecionadas (Figura 20).

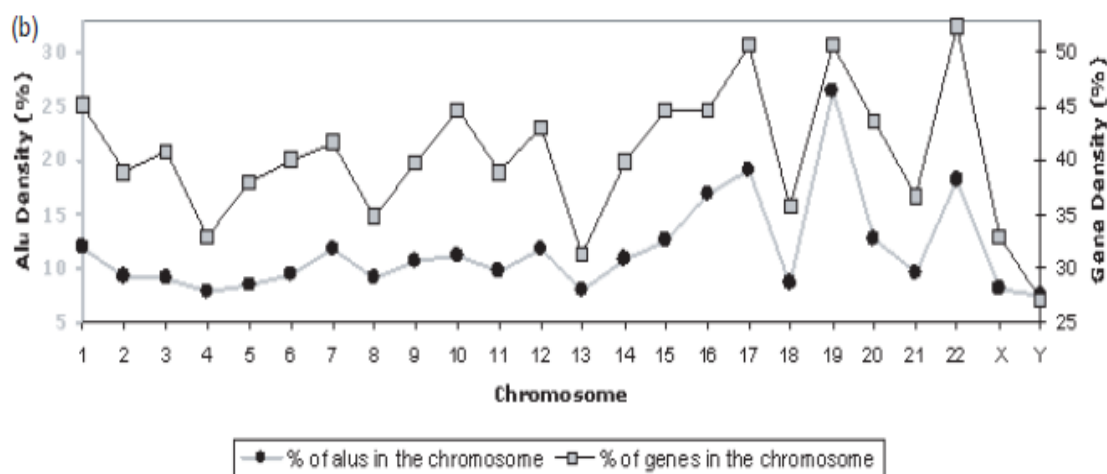


Figura 26. Densidad de secuencias tipo Alu y genes en los distintos cromosomas. Imagen adaptada de Grover *et al.*, 2004.

Además, en ambos genes el número de variantes estructurales es similar al total de variantes que causan una proteína truncada (*nonsense*, *frameshift*). Sin embargo, excepto para los genes *PRPF31*, *CRX*, *TOPORS*, y *RP1*, las variantes tipo *missense* fueron la causa más frecuente de adRP

en nuestra cohorte, mediando un mecanismo etiopatogénico distinto de la haploinsuficiencia de la proteína. Así, el mecanismo asociado a la RP parece ser específico de cada gen, pero también podría estar relacionado con el tipo de defecto y la localización del mismo en el gen (Latasiewicz *et al.*, 2017).

En este trabajo se observó que la adRP presenta un amplio espectro mutacional, al menos en nuestra cohorte, detectándose un total de 114 variantes distintas en las 155 familias caracterizadas. El 46% (53/114) de los cambios identificados en estas familias no habían sido descritos previamente, lo que nos da una idea de que las causas de la adRP en nuestra población son eventos exclusivos de cada familia. Estas nuevas variantes fueron clasificadas como causantes de la enfermedad, basándonos en la elevada conservación de los aminoácidos alterados en proteínas ortólogas, la predicción de su patogenicidad mediante programas "*in silico*", su frecuencia extremadamente rara o ausente en bases de datos poblacionales y, en función de disponibilidad de muestras, su cosegregación familiar con la adRP.

La mutación más frecuente fue p.Pro347Leu, identificada en 10 de las familias representando así el 3,9% en nuestra cohorte, frecuencia similar a la descrita en otras poblaciones europeas (Ziviello *et al.*, 2005; Audo *et al.*, 2010a). Esta variante recurrente afecta a una secuencia CpG con una mayor probabilidad de transición C>T y está considerada como mutación *hotspot*. La segunda mutación más frecuente p.Gly56Arg en *NR2E3*, presente en 7 de las familias, por el contrario parece representar un efecto fundador en nuestra cohorte tal como pone de manifiesto un estudio anterior (Blanco-Kelly *et al.*, 2016). En este sentido, durante esta tesis, hemos investigado la causa de la recurrencia de 2 mutaciones en *PRPF31* en dos familias mediante análisis de haplotipos, lo que ha permitido confirmar un origen ancestral común. Sería interesante la realización de estudios similares en otras variantes recurrentes en nuestra población.

En nuestra cohorte, los genes más prevalentes fueron *RHO*, *PRPF31*, *RP1* y *PRPH2*, explicando el 20,9%, 8,1%, 4,3% y 3,9% de nuestras familias, respectivamente. La tabla 16, muestra el porcentaje de estos cuatro genes en distintas cohortes con adRP. La prevalencia génica en nuestra población fue muy similar a la encontrada en población norteamericana con la excepción del gen *RP1*, que es el quinto gen más prevalente tras *RHO*, *PRPF31*, *PRPH2* y *RPGR* (Daiger *et al.*, 2014a). *RHO* es el gen más frecuentemente mutado en las distintas poblaciones con adRP a nivel mundial sin embargo, la prevalencia de este gen parece cambiar bastante según la población estudiada y la técnica empleada, con prevalencias que varían del 7% en Sudáfrica (Roberts *et al.*, 2000) al 31% en América (Daiger *et al.*, 2014a).

Tabla 16. Porcentaje de caracterización y prevalencia de los genes *RHO*, *PRPF31*, *RP1* y *PRPH2* en distintas poblaciones de adRP. Caract.: caracterización. No: número. *Varias técnicas usadas incluída la NGS.

Población	% Caract.	No. casos de adRP	No. genes estudiados	Técnica	% <i>RHO</i>	% <i>PRPF31</i>	% <i>RP1</i>	% <i>PRPH2</i>	Referencia
Española	60,5	258	-	Varias*	20,9	8,1	4,3	3,9	Presente estudio
Americana	78,5	270	-	Varias*	30,7	8,9	4,8	7	(Daiger <i>et al.</i> , 2014a)
Belga	56	86	-	Varias*	14	10,5	10,5	4,7	(Van Cauwenbergh <i>et al.</i> , 2017)
Franco canadiense	40	60	9	Sanger	18,3	3,3	-	5	(Coussa <i>et al.</i> , 2015)
Japonesa	35,1	77	193	NGS	9	3,9	3,9	5,2	(Oishi <i>et al.</i> , 2014)
Italiana	27	43	12	Sanger	16	-	5	-	(Ziviello <i>et al.</i> , 2005)
China	15	60	<i>RHO</i>	Sanger	15	-	-	-	(Yang <i>et al.</i> , 2014)
Sudáfrica	6,7	60	<i>RHO</i>	SSCP	6,7	-	-	-	(Roberts <i>et al.</i> , 2000)
India	2	53	4	Sanger	-	2	-	-	(Gandra <i>et al.</i> , 2008)

El uso de NGS nos ha permitido identificar variantes en genes raros de adRP, como el gen *HK1*, y determinar que algunos genes considerados minoritarios hasta este momento, presentan una prevalencia mayor a la observada anteriormente en nuestra cohorte (Blanco-Kelly *et al.*, 2012). Un claro ejemplo de ello, en el que profundizaremos más adelante, es el gen *PRPF31*. Por el contrario, aunque el gen *RHO*, sigue siendo el más prevalente en nuestra cohorte, aparece infrarrepresentado en los estudios realizados mediante NGS, debido a su cribado previo mediante secuenciación Sanger en la mayoría de las familias (Fernandez-San Jose *et al.*, 2015b). Los datos de frecuencia de este gen de hecho no han variado con el estudio de 58 nuevas familias con respecto al trabajo realizado en 200 familias anteriormente (Fernandez-San Jose *et al.*, 2015b).

El desarrollo de las técnicas de secuenciación masiva ha venido acompañado de nuevos desafíos en la interpretación de la patogenicidad de las variantes detectadas. A diferencia de las formas recesivas, en las que algunas mutaciones pueden estar enriquecidas en algunas poblaciones, en el caso de adRP las variantes causales son extremadamente raras (Daiger *et al.*, 2014a). Por ello, se adaptó nuestro algoritmo bioinformático de priorización de variantes por frecuencia alélica al análisis específico de formas dominantes utilizando valores de $MAF < 0,001$, lo que permitió disminuir el número de VUS a analizar. Además, el uso de nuevas bases de datos de frecuencias alélicas con datos procedentes de genomas para evitar sesgos procedentes de la captura de los exomas, así como el uso de un mayor número de predictores *in silico*, facilitan la interpretación de variantes raras. Por otro lado, se evaluó su posible implicación de las variantes como responsable de la RP también con su segregación en las familias; sin embargo, estos análisis se ven limitados debido a la penetrancia incompleta, tan frecuentemente asociados a algunas formas de adRP (Martinez-Gimeno *et al.*, 2003; Gamundi *et al.*, 2006), o a la tardía aparición de síntomas, por lo que algunos miembros de la familia pueden no mostrar evidencias de RP en el momento del estudio. Todo ello dificulta la interpretación de la patogenicidad en variantes nuevas, por lo que resulta imprescindible realizar seguimientos de las familias para poder establecer y/o reevaluar la posible causalidad de nuevas variantes.

En nuestro estudio, el análisis de los pacientes con adRP mediante NGS se ha realizado utilizando diferentes paneles génicos que incluían genes no solo asociados a distintas formas de RP sino también a otras formas menos frecuentes de DR. Esta estrategia nos permitió realizar un análisis libre de hipótesis de forma independiente al fenotipo y al modo de herencia de los pacientes en estudio. Gracias a ello, 17 familias previamente clasificadas como formas autosómicas dominantes de RP, fueron reclasificadas clínica y/o genéticamente, debido a la identificación de una variante en un gen asociado a un tipo distinto de adRP.

En 4 familias, encontramos la variante patogénica p.Ser163Arg en el gen *C1QTNF5*, asociado a un fenotipo característico de degeneración retiniana de aparición tardía (LORD: *late-onset retinal degeneration*) (Hayward *et al.*, 2003; Almoguera *et al.*, 2015). Esta forma de DR es una enfermedad ocular con herencia autosómica dominante caracterizada principalmente por ceguera nocturna en la quinta o sexta década de la vida, con moteado amarillo-blanquecino, degeneración central y periférica, y una neovascularización en la coroides en el fondo de ojo (Hayward *et al.*, 2003). Los casos índices de estas cuatro familias presentaban ceguera nocturna como primer síntoma a los 50-60 años y en consecuencia fueron remitidos al Servicio de Genética con un diagnóstico inicial de RP de aparición tardía. Además, alguno de los pacientes presentaba alteración macular y material drusoide en el fondo de ojo. La DR tipo LORD puede ser erróneamente clasificada como una degeneración macular asociada a la edad (DMAE) o como una RP clásica en etapas tempranas o tardías de la enfermedad, respectivamente (Almoguera *et al.*, 2015). A la vista de la relativa frecuencia de esta variante en nuestra cohorte y la dificultad del diagnóstico clínico del LORD, creemos que este gen debe tenerse en cuenta en el estudio de NGS de los pacientes con diagnóstico clínico de RP de aparición tardía.

Además, 8 familias inicialmente clasificadas como adRP, debido a la presencia de mujeres afectas, portaban variantes patogénicas en genes ligados al cromosoma X. Este hallazgo había sido descrito previamente en el 8,5% de las familias con diagnóstico inicial de adRP (Churchill *et al.*, 2013). Aunque las mujeres portadoras de mutaciones en el gen *CHM* son normalmente asintomáticas (Ma *et al.*, 2017), existen descritos en la literatura algunos casos de mujeres completamente afectas de coroideremia (Fraser & Friedmann, 1968). Este hecho se debe a la lionización del cromosoma X no aleatoria, inactivándose el cromosoma X con el alelo *wild-type* (Perez-Cano *et al.*, 2009). La coroideremia es generalmente un diagnóstico clínico realizado a través de hallazgos en el fondo de ojo y evaluación de la historia familiar (MacDonald *et al.*, 1993). En nuestras dos familias con variantes patogénicas en el gen *CHM*, los datos fenotípicos disponibles y los datos del pedigrí eran escasos, por lo tanto, el diagnóstico se basó principalmente en los resultados del estudio molecular.

Por otro lado, se encontraron variantes patogénicas para genes recesivos en 5 familias adicionales. La familia RP-1217 presentó un patrón pseudo-dominante que se explica en el contexto de una elevada endogamia y por la presencia de una variante en homocigosis en el gen *EYS* que segregaba en varias generaciones consecutivas. Además, la herencia pseudo-dominante puede estar justificada también por la frecuencia de portadores relativamente elevada en población española de algunos genes de DR (Riveiro-Alvarez *et al.*, 2009). Este es el caso de la

familia RP-1455, con tres variantes patogénicas diferentes en el gen *ABCA4* segregando en varias generaciones.

Por ello, se pone en evidencia que el uso de un enfoque libre de hipótesis en el análisis de los datos de NGS en pacientes con adRP, permite la reclasificación genética y/o fenotípica de un número importante de familias con sospecha inicial de adRP, lo que tiene un gran impacto en el asesoramiento genético familiar, el seguimiento clínico y la futura inclusión en tratamientos basados en terapia génica.

En la actualidad, se estima que un 40% de los pacientes con DR sigue sin caracterización genética tras análisis con NGS (Audo *et al.*, 2012; Shanks *et al.*, 2013), cifra similar a la observada en nuestro estudio. Varias son las razones que podrían explicar la ausencia de detección de variantes causales en los pacientes estudiados no caracterizados. En primer lugar, algunas familias podrían portar variantes que causan enfermedades en regiones no estudiadas por nuestros paneles de NGS, como en las regiones intrónicas profundas, regiones reguladoras o en regiones altamente repetitivas que son difíciles de capturar con estas técnicas, como sucede con la región *ORF15* en el gen *RPGR*. Además, pueden presentar variantes en genes de adRP recientemente identificados que no estaban incluidos en los paneles usados (como son *ADIPOR1*, *ARL3*, *PRPF4* y *SPP2*) (Chen *et al.*, 2014; Liu *et al.*, 2015; J. Zhang *et al.*, 2016; Strom *et al.*, 2016), o en genes aún no identificados. En segundo lugar, a pesar de la identificación exitosa de reordenamientos genómicos en este estudio, no podemos descartar la posibilidad de que existan CNVs adicionales que involucren regiones parcialmente cubiertas por los enfoques de NGS utilizados.

En el caso de las variantes no codificantes, los diseños de paneles dirigidos no suelen incluir las regiones intrónicas profundas de los genes diana debido a limitaciones de cobertura. Sin embargo, se han descrito variantes intrónicas en distintos genes asociados a DR como *USH2A* (Garcia-Garcia *et al.*, 2014), *ABCA4* (Zernant *et al.*, 2014) y *OFD1* (Webb *et al.*, 2012), entre otros, que conducen a la inclusión de pseudoexones mediante la activación de sitios crípticos de *splicing*. En nuestro panel se incluyeron algunas de las regiones intrónicas que habían sido implicadas en DR en el momento del diseño (Perez-Carro *et al.*, 2016). Cada vez existen más evidencias de la implicación de variantes sinónimas en distintas patologías (Hunt *et al.*, 2014) conduciendo a desregulación del *splicing*, sin embargo, éstas no suelen ser consideradas durante el filtrado de variantes. Además, tampoco se suelen analizar otros reguladores como *miRNAs*, que también pueden estar implicados en algunas formas de DR, como ya ha sido descrito previamente para el *mir-204* en una familia dominante con DR y coloboma bilateral (Conte *et*

al., 2015). En un futuro, el análisis de estas regiones no codificantes en genes asociados a adRP podría aumentar el número de pacientes caracterizados.

5.2. Algoritmo diagnóstico actual para las familias con adRP

La combinación de diferentes estrategias moleculares en la adRP a lo largo del tiempo, nos ha permitido alcanzar y mejorar la alta tasa diagnóstica en nuestra cohorte española. Nuestro estudio confirma el gran valor diagnóstico de las herramientas de NGS basadas en paneles de genes, que representa actualmente el enfoque de mayor confianza, rentabilidad y eficacia para la identificación no solo de SNVs causantes de enfermedades, sino también de CNVs en familias con adRP. Por lo tanto, dejamos el uso de WES y/o WGS para los pacientes no caracterizados a nivel de investigación, ya que podría ayudarnos a identificar nuevos genes asociados a adRP y así, poder caracterizar mayor número de familias (Figura 27).

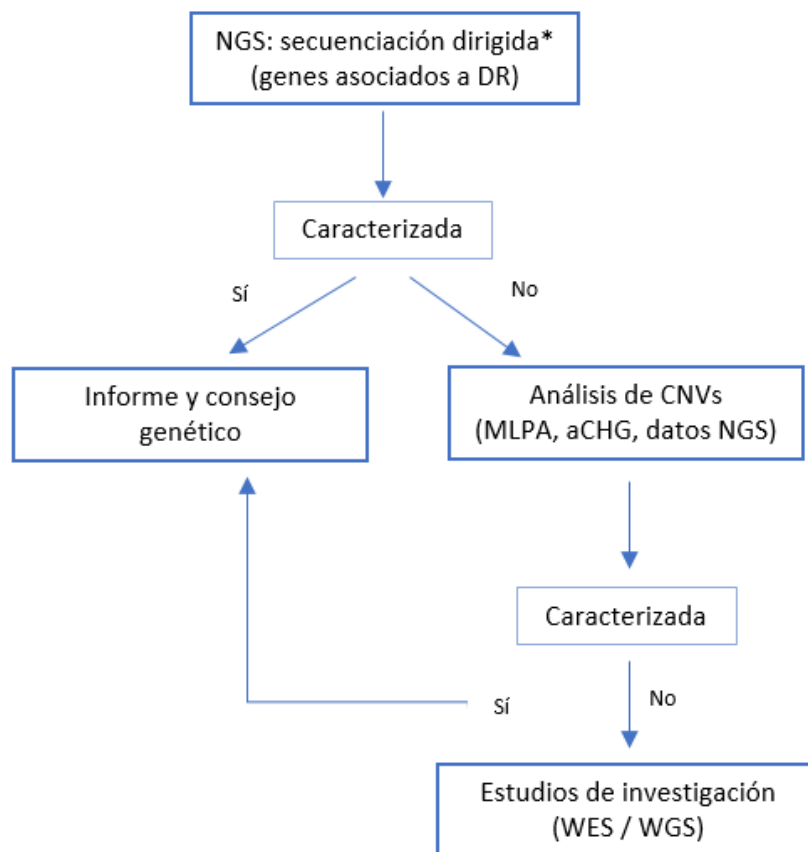


Figura 27. Algoritmo diagnóstico propuesto para las familias con adRP. * paneles físicos o virtuales a partir de Exoma clínico.

5.3. Importancia del gen *PRPF31* en la adRP

Nuestro trabajo puso en evidencia que las mutaciones en el gen *PRPF31* explican el 8,1% de las familias, lo que supone una frecuencia superior a la descrita previamente en otro estudio en población española (1,7%) (Martinez-Gimeno *et al.*, 2003). Este dato es similar a lo descrito en población americana (8,9%) (Daiger *et al.*, 2014a) y superior a la observada en las poblaciones francesa (Audo *et al.*, 2010b), británica (Waseem *et al.*, 2007), japonesa (Sato *et al.*, 2005) y china (Lim *et al.*, 2009), con frecuencias de 6,7%, 5%, 3% y 1%, respectivamente. La frecuencia de reordenamientos genómicos en el gen *PRPF31* en población española es del 1,9%. Esta prevalencia es inferior a la descrita previamente en una población de América del Norte (Sullivan *et al.*, 2006b), pero más alta que en Francia (Audo *et al.*, 2010b), donde no se encontraron reordenamientos genómicos.

El espectro mutacional en el gen *PRPF31* observado en nuestra cohorte es muy heterogéneo, encontrándose un total de 19 mutaciones distintas en 20 familias diferentes estudiadas. Además, describimos 4 nuevas mutaciones puntuales, 3 deleciones grandes y, por primera vez, una duplicación de varios exones asociada a adRP. No se encontraron puntos calientes o mutaciones recurrentes en este gen, salvo dos variantes. La variante p.Glu181*, descrita previamente en otra familia española (Pomares *et al.*, 2010), fue detectada en dos familias con mismo origen geográfico, que compartían una región común de 1,2Mb, lo que sugiere un efecto fundador. De forma similar, la variante c.322+4_322+7del, descrita por primera vez (Q. Zhang *et al.*, 2016) como mutación fundadora en Cuba, está presente en dos de nuestras familias de distinto origen (Cuba y Madrid), que comparten una región común de mínimo 1,7Mb, sugiriendo un efecto fundador probablemente de origen hispánico, respaldado por la historia de la migración de población desde España a América desde el siglo XVI (Q. Zhang *et al.*, 2016).

Adicionalmente, en la familia RP-0932 se encontraron dos grandes reordenamientos distintos segregando en dos ramas diferentes de la misma familia. Se han descrito grandes deleciones en el gen *PRPF31* que incluyen varios exones contiguos, el gen completo o incluso parte del gen junto a otros genes aguas arriba del gen (Sullivan *et al.*, 2006b; Eisenberger *et al.*, 2013), como lo que hemos visto en nuestra cohorte. Pero esta es la primera vez que dos tipos de reordenamientos en *PRPF31* (una duplicación y una deleción), segregan en diferentes individuos de una misma familia con adRP.

La mayoría de las mutaciones encontradas en este gen en nuestra cohorte fueron mutaciones de pérdida de función (LOF: *loss-of-function*) por ejemplo, mutaciones tipo *nonsense*, de *splicing*, *frameshift* y grandes reordenamientos, este tipo de variantes provocarán un codón de parada prematuro, siendo el transcrito susceptible de ser degradado vía NMD (*nonsense-mediated*

mRNA decay) (Rio Frio *et al.*, 2008). Por ello, la mayoría de las mutaciones en *PRPF31* parecen causar enfermedad por haploinsuficiencia (Rio Frio *et al.*, 2008). En la familia RP-0932, la expresión del gen *PRPF31* se redujo de manera significativa del 30 al 48%, tanto en los pacientes con la delección como en los pacientes con la duplicación, respecto a los niveles de expresión en las muestras control. Este hallazgo apoya la hipótesis de la haploinsuficiencia como mecanismo patogénico asociado al gen *PRPF31*. Por ello, para el caso de la duplicación en tándem, la causa más probable de haploinsuficiencia es la degradación de los transcritos aberrantes del gen vía NMD. Mientras que en los pacientes con delecciones en *PRPF31* con la región 5'UTR en el exón 1 no codificante delecionada, el cual se predice que contiene varios sitios de regulación transcripcional, y por lo tanto la haploinsuficiencia puede explicarse por una desregulación de la transcripción génica.

A la vista de la elevada heterogeneidad alélica y complejidad de los mecanismos mutacionales asociados a *PRPF31*, los estudios moleculares para adRP requieren no solo la secuenciación de los exones y los sitios de *splicing* de este gen, sino también un análisis de CNV para detectar grandes reordenamientos mediante análisis de datos de NGS, MLPA o aCGH.

La penetrancia incompleta (PI) es otra de las características del gen *PRPF31* (Daiger *et al.*, 2014a), observándose en un gran número de familias con mutaciones en *PRPF31*, concretamente en el 65% de las familias de nuestra cohorte. El mecanismo que explica este fenómeno de la PI es un desequilibrio alélico con un aumento de la expresión del alelo *wild-type*, compensando los niveles de expresión del alelo mutado en los pacientes portadores asintomáticos (Vithana *et al.*, 2003; Rivolta *et al.*, 2006). Esta sobreexpresión del gen (Venturini *et al.*, 2012; Rose *et al.*, 2014; Rose *et al.* 2016), parece ser debida a factores genéticos modificadores de la expresión de *PRPF31*, como el gen *CNOT* o el elemento repetido MSR1 aguas arriba del promotor del gen (Rose & Bhattacharya, 2016), que están implicados en la represión transcripcional de *PRPF31* (Venturini *et al.*, 2012; Rose *et al.*, 2014; Rose *et al.*, 2016). Sin embargo, el mecanismo exacto de la PI no está aún claro.

La presencia de 3 o 4 copias del minisatélite MSR1 mostró fuertes diferencias en la modulación de la transcripción del gen *PRPF31*, que podría estar asociado a la PI en la RP asociada a mutaciones en *PRPF31* (Rose *et al.*, 2016). De acuerdo con la hipótesis de que la presencia de 4 copias de MSR1 parece ser un importante factor protector para la expresividad de la enfermedad (Rose *et al.*, 2016), los individuos asintomáticos de nuestra cohorte presentaban una frecuencia del alelo de 4 copias significativamente superior al de la población control. Rose y colaboradores reportaron que todos los pacientes sintomáticos en su estudio, eran

homocigotos para el alelo de 3 copias, prediciendo tener una expresión disminuida del gen y desarrollando por tanto, la enfermedad (Rose *et al.*, 2016). Sin embargo, el 33% de nuestras familias (RP-0932, RP-0368, RP-0637, RP-0361, y RP-1599), presentaron al menos un individuo sintomático heterocigoto 3/4. Este hecho es contradictorio con lo descrito por Rose y colaboradores, reflejando una mayor frecuencia del alelo 4 en los individuos sintomáticos en nuestra población, que la reportada previamente (Rose *et al.*, 2016). En estos casos heterocigóticos sintomáticos, se espera que la expresión del gen *PRPF31* esté reducida, desarrollándose los síntomas, si el alelo *wild-type* se encuentra en *cis* con el alelo MSR1 de 3 repeticiones y el alelo 4 MSR1 está en *cis* con el alelo mutado, no actuando el alelo de 4 copias como factor protector. Por el contrario, en los individuos heterocigóticos asintomáticos, el alelo 4 protector debería estar en *cis* con el alelo *wild-type* para mantener una alta expresión de *PRPF31* y así compensar la expresión del alelo mutado.

Desafortunadamente, no fue posible confirmar completamente esta hipótesis en nuestro estudio, dado que la fase completa de los alelos no se pudo inferir en la mayoría de las familias debido a la falta de información familiar y/o disponibilidad de muestras. Sin embargo, observamos dos excepciones claras para esta hipótesis en dos individuos heterocigotos 3/4 asintomáticos, los individuos III:6 y III:9 de las familias RP-0361 y RP-1599, respectivamente, donde el alelo 4 MSR1 está necesariamente en *cis* con el alelo mutado, como demuestra el análisis de haplotipos (Anexo 3). En estos dos casos, MSR1 no puede explicar el estado asintomático, por lo que debe existir algún otro factor transactivador del gen *PRPF31* que permita explicar la PI en estas familias. Esta última observación está de acuerdo con la hipótesis de un control poligénico de la regulación transcripcional del gen *PRPF31* para explicar la PI de la RP asociada a *PRPF31* (Rose & Bhattacharya, 2016).

Además, se realizó una revisión oftalmológica en 31 pacientes de 18 familias diferentes que portaban mutaciones en *PRPF31*, que representa la descripción fenotípica más grande para este gen en la actualidad. Los signos clínicos observados en la familia RP-0932 parecen ser más severos en los pacientes portadores de la duplicación en comparación con los pacientes portadores de la delección en *PRPF31*. Por lo tanto, esta familia es un buen ejemplo de variabilidad fenotípica asociada a diferentes defectos en *PRPF31*. En nuestra cohorte, los pacientes portadores de una mutación tipo *missense*, en comparación con aquellos pacientes portadores de una mutación *nonsense* o *frameshift*, no presentaron necesariamente un fenotipo más leve como describió Xu y colaboradores (Xu *et al.*, 2012). Aunque estos resultados son preliminares y no pueden usarse para establecer una clara correlación genotipo-fenotipo porque la información clínica para pacientes portadores de defectos en *PRPF31* fue limitada en

algunos casos, parece que la duplicación y las mutaciones tipo *missense* podrían producir un fenotipo más severo que el resto de mutaciones. Los pacientes de las familias RP-0932 y RP-0777 con delección en genes adicionales aguas arriba del gen *PRPF31*, no mostraron hallazgos clínicos adicionales aparte de la RP, lo cual es consistente con las observaciones previas (Abu-Safieh *et al.*, 2006; Kohn *et al.*, 2009; Rose *et al.*, 2011).

A la luz de la alta heterogeneidad genética de las mutaciones en *PRPF31* es necesario estudiar todos los exones y las primeras bases de los intrones implicados en el *splicing*, además de realizar un análisis de CNVs, debido a la alta cantidad de grandes duplicaciones o deleciones encontradas en el gen. La identificación de nuevos mecanismos patogénicos subyacentes a la RP por mutaciones en *PRPF31* y la correlación genotipo-fenotipo no solo son importantes para el diagnóstico actual sino también para el asesoramiento genético. Además, es particularmente importante para el desarrollo de nuevas terapias génicas y una futura inclusión de los pacientes en ensayos clínicos.

Conclusiones

- 1.- El uso de las nuevas técnicas de secuenciación masiva ha permitido aumentar la tasa de caracterización molecular de la retinosis pigmentaria autosómica dominante en población española, desde el 39% con las técnicas clásicas, hasta la tasa diagnóstica actual del 60%.
- 2.- Los genes más prevalentes encontrados en nuestra población de adRP son *RHO* (20,9%), *PRPF31* (8,1%), *RP1* (4,3%) y *PRPH2* (3,9%). Estos cuatro genes constituyen el 61,9% (96/155) de las familias resueltas.
- 3.- El tipo de mutación encontrada en los genes asociados a adRP son de tipo *missense*, salvo para los genes *PRPF31*, *CRX*, *TOPORS* y *RP1*, en los que se observan mutaciones tipo LOF.
- 4.- El análisis de genes de DR mediante NGS –asociados a distintos patrones de herencia y fenotipos- de las familias con sospecha de adRP, nos ha permitido caracterizar y reclasificar genética y clínicamente a un 6,6% de las familias, presentando 3,1% de las familias genes ligados al cromosoma X, 1,9% genes recesivos y 1,6% presentaron la variante p.Ser163Arg en el gen *C1QTNF5*. Ello ha proporcionado a estas familias un adecuado consejo genético y su posible inclusión en futuras aproximaciones terapéuticas.
- 5.- El análisis de CNVs nos ha permitido caracterizar a un 3% de familias con sospecha de adRP. El 1,9% (5/258) de las alteraciones se encontraron en el gen *PRPF31*, seguidos del gen *CRX* (0,8%; 2/258) y *RHO* (0,4%, 1/258).
- 6.- Como algoritmo para la caracterización molecular de los casos de adRP se propone la aplicación de secuenciación masiva dirigida como primer abordaje diagnóstico y posterior análisis de variaciones en el número de copia. Y en una segunda fase, ya en el marco de la investigación, el uso de la secuenciación del exoma completo o el genoma en casos negativos, para identificar nuevos genes asociados a distrofias de retina.
- 7.- El estudio de *PRPF31*, segundo gen más prevalente en familias españolas con adRP, ha aportado información de la elevada tasa de CNVs en este gen, siendo imprescindible realizar un análisis de CNVs en familias con adRP, especialmente en familias con salto de generaciones en los afectos, debido a la elevada tasa (65%) de penetrancia incompleta en las familias con mutaciones en este gen.
- 8.- El estudio de CNVs en el gen *PRPF31* permitió encontrar una duplicación y una delección segregando en diferentes individuos de la misma familia. El análisis de expresión de los pacientes

tanto con duplicación como con delección, con disminución de la expresión del gen respecto a los controles en un 30% al 48%, apoyan la haploinsuficiencia como mecanismo patogénico.

9.- El análisis del número de copias del minisatélite MSR1 en las familias con mutaciones en *PRPF31*, permitió ver que los portadores asintomáticos presentaban un aumento significativo del alelo 4 respecto a los controles, pero mediante el análisis de haplotipos en 2 de las familias (RP-0361 y RP-1599) el alelo 4 se encontró en *cis* con el alelo mutado, no pudiendo actuar como factor protector. Por ello, al menos en estas familias, debe de existir otro u otros agentes moduladores de la expresión de este gen.

10.- Tras la correlación genotipo-fenotipo del gen *PRPF31*, se confirmó que la ceguera nocturna es el primer síntoma, seguido de constricción del campo visual y pérdida de la agudeza visual. Los síntomas más severos se observaron en los pacientes con la duplicación y las mutaciones *missense*, en términos de edad de aparición y curso de la enfermedad, aunque no se puede obtener una correlación general de genotipo-fenotipo debido al bajo número de pacientes estudiados con estos cambios.

Bibliografia

A

- Abu-Safieh, L., Vithana, E. N., Mantel, I., Holder, G. E., Pelosini, L., et al. (2006) A large deletion in the adRP gene PRPF31: evidence that haploinsufficiency is the cause of disease. *Mol Vis* 12: 384-388.
- Al-Magthteh, M., Gregory, C., Inglehearn, C., Hardcastle, A. & Bhattacharya, S. (1993) Rhodopsin mutations in autosomal dominant retinitis pigmentosa. *Hum Mutat* 2(4): 249-255.
- Al-Magthteh, M., Vithana, E., Tarttelin, E., Jay, M., Evans, K., et al. (1996) Evidence for a major retinitis pigmentosa locus on 19q13.4 (RP11) and association with a unique bimodal expressivity phenotype. *Am J Hum Genet* 59(4): 864-871.
- Ali, M. U., Rahman, M. S. U., Cao, J. & Yuan, P. X. (2017) Genetic characterization and disease mechanism of retinitis pigmentosa; current scenario. *3 Biotech* 7(4): 251.
- Almoguera, B., Li, J., Fernandez-San Jose, P., Liu, Y., March, M., et al. (2015) Application of Whole Exome Sequencing in Six Families with an Initial Diagnosis of Autosomal Dominant Retinitis Pigmentosa: Lessons Learned. *PLoS One* 10(7): e0133624.
- Andreasson, S., Ehinger, B., Abrahamson, M. & Fex, G. (1992) A six-generation family with autosomal dominant retinitis pigmentosa and a rhodopsin gene mutation (arginine-135-leucine). *Ophthalmic Paediatr Genet* 13(3): 145-153.
- Arshavsky, V. Y., Lamb, T. D. & Pugh, E. N., Jr. (2002) G proteins and phototransduction. *Annu Rev Physiol* 64: 153-187.
- Audo, I., Manes, G., Mohand-Said, S., Friedrich, A., Lancelot, M. E., et al. (2010a) Spectrum of rhodopsin mutations in French autosomal dominant rod-cone dystrophy patients. *Invest Ophthalmol Vis Sci* 51(7): 3687-3700.
- Audo, I., Bujakowska, K., Mohand-Said, S., Lancelot, M. E., Moskova-Doumanova, V., et al. (2010b) Prevalence and novelty of PRPF31 mutations in French autosomal dominant rod-cone dystrophy patients and a review of published reports. *BMC Med Genet* 11: 145.
- Audo, I., Bujakowska, K. M., Leveillard, T., Mohand-Said, S., Lancelot, M. E., et al. (2012) Development and application of a next-generation-sequencing (NGS) approach to detect known and novel gene defects underlying retinal diseases. *Orphanet J Rare Dis* 7: 8.
- Avila-Fernandez, A., Perez-Carro, R., Corton, M., Lopez-Molina, M. I., Campello, L., et al. (2016) Whole-exome sequencing reveals ZNF408 as a new gene associated with autosomal recessive retinitis pigmentosa with vitreal alterations. *Hum Mol Genet* 24(14): 4037-4048.
- Ayuso, C., Garcia-Sandoval, B., Najera, C., Valverde, D., Carballo, M. & Antinolo, G. (1995) Retinitis pigmentosa in Spain. The Spanish Multicentric and Multidisciplinary Group for Research into Retinitis Pigmentosa. *Clin Genet* 48(3): 120-122.

Ayuso, C. & Millan, J. M. (2010) Retinitis pigmentosa and allied conditions today: a paradigm of translational research. *Genome Med* 2(5): 34.

B

Bao, R., Huang, L., Andrade, J., Tan, W., Kibbe, W. A., et al. (2014) Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform* 13(Suppl 2): 67-82.

Benaglio, P., McGee, T. L., Capelli, L. P., Harper, S., Berson, E. L. & Rivolta, C. (2011) Next generation sequencing of pooled samples reveals new SNRNP200 mutations associated with retinitis pigmentosa. *Hum Mutat* 32(6): E2246-2258.

Bernardis, I., Chiesi, L., Tenedini, E., Artuso, L., Percesepe, A., et al. (2016) Unravelling the Complexity of Inherited Retinal Dystrophies Molecular Testing: Added Value of Targeted Next-Generation Sequencing. *Biomed Res Int* 2016: 6341870.

Bhavsar, P. K., Brand, N. J., Yacoub, M. H. & Barton, P. J. (1996) Isolation and characterization of the human cardiac troponin I gene (TNNI3). *Genomics* 35(1): 11-23.

Blanco-Kelly, F., Garcia-Hoyos, M., Corton, M., Avila-Fernandez, A., Riveiro-Alvarez, R., et al. (2012) Genotyping microarray: mutation screening in Spanish families with autosomal dominant retinitis pigmentosa. *Mol Vis* 18: 1478-1483.

Blanco-Kelly, F., Garcia Hoyos, M., Lopez Martinez, M. A., Lopez-Molina, M. I., Riveiro-Alvarez, R., et al. (2016) Dominant Retinitis Pigmentosa, p.Gly56Arg Mutation in NR2E3: Phenotype in a Large Cohort of 24 Cases. *PLoS One* 11(2): e0149473.

Borras, E., de Sousa Dias, M., Hernan, I., Pascual, B., Mane, B., et al. (2013) Detection of novel genetic variation in autosomal dominant retinitis pigmentosa. *Clin Genet* 84(5): 441-452.

Bowne, S. J., Daiger, S. P., Hims, M. M., Sohocki, M. M., Malone, K. A., et al. (1999) Mutations in the RP1 gene causing autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 8(11): 2121-2128.

Bowne, S. J., Sullivan, L. S., Blanton, S. H., Cepko, C. L., Blackshaw, S., et al. (2002) Mutations in the inosine monophosphate dehydrogenase 1 gene (IMPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 11(5): 559-568.

Bowne, S. J., Sullivan, L. S., Koboldt, D. C., Ding, L., Fulton, R., et al. (2011) Identification of disease-causing mutations in autosomal dominant retinitis pigmentosa (adRP) using next-generation DNA sequencing. *Invest Ophthalmol Vis Sci* 52(1): 494-503.

Bravo-Gil, N., Mendez-Vidal, C., Romero-Perez, L., Gonzalez-del Pozo, M., Rodriguez-de la Rua, E., et al. (2016) Improving the management of Inherited Retinal Dystrophies by targeted sequencing of a population-specific gene panel. *Sci Rep* 6: 23910.

Bujakowska, K. M., Fernandez-Godino, R., Place, E., Consugar, M., Navarro-Gomez, D., et al. (2017) Copy-number variation is an important contributor to the genetic causality of inherited retinal degenerations. *Genet Med* 19(6): 643-651.

C

Cao, H., Wu, J., Lam, S., Duan, R., Newnham, C., et al. (2011) Temporal and tissue specific regulation of RP-associated splicing factor genes PRPF3, PRPF31 and PRPF38--implications in the pathogenesis of RP. *PLoS One* 6(1): e15860.

Chakarova, C. F., Hims, M. M., Bolz, H., Abu-Safieh, L., Patel, R. J., et al. (2002) Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 11(1): 87-92.

Chakarova, C. F., Cherninkova, S., Tournev, I., Waseem, N., Kaneva, R., et al. (2006) Molecular genetics of retinitis pigmentosa in two Romani (Gypsy) families. *Mol Vis* 12: 909-914.

Chen, X., Liu, Y., Sheng, X., Tam, P. O., Zhao, K., et al. (2014) PRPF4 mutations cause autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 23(11): 2926-2939.

Churchill, J. D., Bowne, S. J., Sullivan, L. S., Lewis, R. A., Wheaton, D. K., et al. (2013) Mutations in the X-linked retinitis pigmentosa genes RPGR and RP2 found in 8.5% of families with a provisional diagnosis of autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 54(2): 1411-1416.

Conte, I., Hadfield, K. D., Barbato, S., Carrella, S., Pizzo, M., et al. (2015) MiR-204 is responsible for inherited retinal dystrophy associated with ocular coloboma. *Proc Natl Acad Sci U S A* 112(25): E3236-3245.

Coppieters, F., Leroy, B. P., Beysen, D., Hellemans, J., De Bosscher, K., et al. (2007) Recurrent mutation in the first zinc finger of the orphan nuclear receptor NR2E3 causes autosomal dominant retinitis pigmentosa. *Am J Hum Genet* 81(1): 147-157.

Corton, M., Nishiguchi, K. M., Avila-Fernandez, A., Nikopoulos, K., Riveiro-Alvarez, R., et al. (2013) Exome sequencing of index patients with retinal dystrophies as a tool for molecular diagnosis. *PLoS One* 8(6): e65574.

Corton, M., Avila-Fernandez, A., Campello, L., Sanchez, M., Benavides, B., et al. (2016) Identification of the Photoreceptor Transcriptional Co-Repressor SAMD11 as Novel Cause of Autosomal Recessive Retinitis Pigmentosa. *Sci Rep* 6: 35370.

Coussa, R. G., Chakarova, C., Ajlan, R., Taha, M., Kavalec, C., et al. (2015) Genotype and Phenotype Studies in Autosomal Dominant Retinitis Pigmentosa (adRP) of the French Canadian Founder Population. *Invest Ophthalmol Vis Sci* 56(13): 8297-8305.

D

Daiger, S. P., Bowne, S. J. & Sullivan, L. S. (2014a) Genes and Mutations Causing Autosomal Dominant Retinitis Pigmentosa. *Cold Spring Harb Perspect Med* 5(10).

- Daiger, S. P., Bowne, S. J., Sullivan, L. S., Blanton, S. H., Weinstock, G. M., et al. (2014b) Application of next-generation sequencing to identify genes and mutations causing autosomal dominant retinitis pigmentosa (adRP). *Adv Exp Med Biol* 801: 123-129.
- Davidson, A. E., Millar, I. D., Urquhart, J. E., Burgess-Mullan, R., Shweikh, Y., et al. (2009) Missense mutations in a retinal pigment epithelium protein, bestrophin-1, cause retinitis pigmentosa. *Am J Hum Genet* 85(5): 581-592.
- Deininger, P. L. & Batzer, M. A. (1999) Alu repeats and human disease. *Mol Genet Metab* 67(3): 183-193.
- de Sousa Dias, M., Hernan, I., Pascual, B., Borrás, E., Mane, B., et al. (2013) Detection of novel mutations that cause autosomal dominant retinitis pigmentosa in candidate genes by long-range PCR amplification and next-generation sequencing. *Mol Vis* 19: 654-664.
- de Sousa Dias, M., Hernan, I., Delas, B., Pascual, B., Borrás, E., et al. (2015) New COL6A6 variant detected by whole-exome sequencing is linked to break points in intron 4 and 3'-UTR, deleting exon 5 of RHO, and causing adRP. *Mol Vis* 21: 857-870.
- den Hollander, A. I., Black, A., Bennett, J. & Cremers, F. P. (2010) Lighting a candle in the dark: advances in genetics and gene therapy of recessive retinal dystrophies. *J Clin Invest* 120(9): 3042-3053.
- Dryja, T. P., McGee, T. L., Hahn, L. B., Cowley, G. S., Olsson, J. E., et al. (1990) Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa. *N Engl J Med* 323(19): 1302-1307.
- Dryja, T. P., Hahn, L. B., Cowley, G. S., McGee, T. L. & Berson, E. L. (1991) Mutation spectrum of the rhodopsin gene among patients with autosomal dominant retinitis pigmentosa. *Proc Natl Acad Sci U S A* 88(20): 9370-9374.

E

- Eisenberger, T., Neuhaus, C., Khan, A. O., Decker, C., Preising, M. N., et al. (2013) Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: the example of retinal dystrophies. *PLoS One* 8(11): e78496.
- Ellingford, J. M., Horn, B., Campbell, C., Arno, G., Barton, S., et al. (2017) Assessment of the incorporation of CNV surveillance into gene panel next-generation sequencing testing for inherited retinal diseases. *J Med Genet*.
- Eudy, J. D., Weston, M. D., Yao, S., Hoover, D. M., Rehm, H. L., et al. (1998) Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. *Science* 280(5370): 1753-1757.

F

- Farrar, G. J., Kenna, P., Jordan, S. A., Kumar-Singh, R., Humphries, M. M., et al. (1992) Autosomal dominant retinitis pigmentosa: a novel mutation at the peripherin/RDS locus in the original 6p-linked pedigree. *Genomics* 14(3): 805-807.
- Farrar, G. J., Carrigan, M., Dockery, A., Millington-Ward, S., Palfi, A., et al. (2017) Toward an elucidation of the molecular genetics of inherited retinal degenerations. *Hum Mol Genet* 26(R1): R2-R11.
- Fernandez-San Jose, P., Corton, M., Blanco-Kelly, F., Avila-Fernandez, A., Lopez-Martinez, M. A., et al. (2015a) Targeted Next-Generation Sequencing Improves the Diagnosis of Autosomal Dominant Retinitis Pigmentosa in Spanish Patients. *Invest Ophthalmol Vis Sci* 56(4): 2173-2182.
- Fernandez-San Jose, P., Blanco-Kelly, F., Corton, M., Trujillo-Tiebas, M. J., Gimenez, A., et al. (2015b) Prevalence of Rhodopsin mutations in autosomal dominant Retinitis Pigmentosa in Spain: clinical and analytical review in 200 families. *Acta Ophthalmol* 93(1): e38-44.
- Fernández-San-José, P. (2017). *Caracterización clínica y genética de familias españolas con retinosis pigmentaria autosómica dominante mediante secuenciación masiva y otras técnicas de diagnóstico molecular* (Tesis Doctoral). Universidad Autónoma de Madrid, Madrid.
- Ferrari, S., Di Iorio, E., Barbaro, V., Ponzin, D., Sorrentino, F. S. & Parmeggiani, F. (2011) Retinitis pigmentosa: genes and disease mechanisms. *Curr Genomics* 12(4): 238-249.
- Fishman, G. A., Stone, E., Gilbert, L. D., Vandenburg, K., Sheffield, V. C. & Heckenlively, J. R. (1994) Clinical features of a previously undescribed codon 216 (proline to serine) mutation in the peripherin/retinal degeneration slow gene in autosomal dominant retinitis pigmentosa. *Ophthalmology* 101(8): 1409-1421.
- Fraser, G. R. & Friedmann, A. I. (1968) Choroideremia in a female. *Br Med J* 2(5607): 732-734.

G

- Gamundi, M. J., Hernan, I., Martinez-Gimeno, M., Maseras, M., Garcia-Sandoval, B., et al. (2006) Three novel and the common Arg677Ter RP1 protein truncating mutations causing autosomal dominant retinitis pigmentosa in a Spanish population. *BMC Med Genet* 7: 35.
- Gandra, M., Anandula, V., Authiappan, V., Sundaramurthy, S., Raman, R., et al. (2008) Retinitis pigmentosa: mutation analysis of RHO, PRPF31, RP1, and IMPDH1 genes in patients from India. *Mol Vis* 14: 1105-1113.
- Garcia-Garcia, G., Aller, E., Jaijo, T., Aparisi, M. J., Larrieu, L., et al. (2014) Novel deletions involving the USH2A gene in patients with Usher syndrome and retinitis pigmentosa. *Mol Vis* 20: 1398-1410.

- Ge, Z., Bowles, K., Goetz, K., Scholl, H. P., Wang, F., et al. (2015) NGS-based Molecular diagnosis of 105 eyeGENE((R)) probands with Retinitis Pigmentosa. *Sci Rep* 5: 18287.
- Geller, A. M. & Sieving, P. A. (1993) Assessment of foveal cone photoreceptors in Stargardt's macular dystrophy using a small dot detection task. *Vision Res* 33(11): 1509-1524.
- Glockle, N., Kohl, S., Mohr, J., Scheurenbrand, T., Sprecher, A., et al. (2014) Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. *Eur J Hum Genet* 22(1): 99-104.
- Gonzalez-del Pozo, M., Borrego, S., Barragan, I., Pieras, J. I., Santoyo, J., et al. (2011) Mutation screening of multiple genes in Spanish patients with autosomal recessive retinitis pigmentosa by targeted resequencing. *PLoS One* 6(12): e27894.
- Grover, S., Fishman, G. A., Anderson, R. J., Tozatti, M. S., Heckenlively, J. R., et al. (1999) Visual acuity impairment in patients with retinitis pigmentosa at age 45 years or older. *Ophthalmology* 106(9): 1780-1785.
- Grover, D., Mukerji, M., Bhatnagar, P., Kannan, K. & Brahmachari, S. K. (2004) Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. *Bioinformatics* 20(6): 813-817.
- Grupo de trabajo de la Guía de Práctica Clínica para las Distrofias Hereditarias de Retina. Guía de Práctica Clínica para las Distrofias Hereditarias de Retina. Ministerio de Sanidad, Servicios Sociales e Igualdad. Servicio de Evaluación del Servicio Canario de la Salud; 2017. Guías de Práctica Clínica en el SNS.

H

- Hamel, C. (2006) Retinitis pigmentosa. *Orphanet J Rare Dis* 1: 40.
- Hartong, D. T., Berson, E. L. & Dryja, T. P. (2006) Retinitis pigmentosa. *Lancet* 368(9549): 1795-1809.
- Hayakawa, M., Fujiki, K., Hotta, Y., Ito, R., Ohki, J., et al. (1999) Visual impairment and REP-1 gene mutations in Japanese choroideremia patients. *Ophthalmic Genet* 20(2): 107-115.
- Hayward, C., Shu, X., Cideciyan, A. V., Lennon, A., Barran, P., et al. (2003) Mutation in a short-chain collagen gene, CTRP5, results in extracellular deposit formation in late-onset retinal degeneration: a genetic model for age-related macular degeneration. *Hum Mol Genet* 12(20): 2657-2667.
- Hernan, I., Gamundi, M. J., Borrás, E., Maseras, M., Garcia-Sandoval, B., et al. (2012) Novel p.M96T variant of NRL and shRNA-based suppression and replacement of NRL mutants associated with autosomal dominant retinitis pigmentosa. *Clin Genet* 82(5): 446-452.
- Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S. H. & Elledge, S. J. (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev* 23(7): 837-848.

Hunt, R. C., Simhadri, V. L., Iandoli, M., Sauna, Z. E. & Kimchi-Sarfaty, C. (2014) Exposing synonymous mutations. *Trends Genet* 30(7): 308-321.

I

Inglehearn, C. F., Keen, T. J., Bashir, R., Jay, M., Fitzke, F., et al. (1992) A completed screen for mutations of the rhodopsin gene in a panel of patients with autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 1(1): 41-45.

J

Johansson, L. F., van Dijk, F., de Boer, E. N., van Dijk-Bos, K. K., Jongbloed, J. D., et al. (2016) CoNVaDING: Single Exon Variation Detection in Targeted NGS Data. *Hum Mutat* 37(5): 457-464.

K

Kennan, A., Aherne, A., Palfi, A., Humphries, M., McKee, A., et al. (2002) Identification of an IMPDH1 mutation in autosomal dominant retinitis pigmentosa (RP10) revealed following comparative microarray analysis of transcripts derived from retinas of wild-type and Rho(-/-) mice. *Hum Mol Genet* 11(5): 547-557.

Kitiratschky, V. B., Glockner, C. J. & Kohl, S. (2011) Mutation screening of the GUCA1B gene in patients with autosomal dominant cone and cone rod dystrophy. *Ophthalmic Genet* 32(3): 151-155.

Kohl, S., Christ-Adler, M., Apfelstedt-Sylla, E., Kellner, U., Eckstein, A., et al. (1997) RDS/peripherin gene mutations are frequent causes of central retinal dystrophies. *J Med Genet* 34(8): 620-626.

Kohn, L., Bowne, S. J., L, S. S., Daiger, S. P., Burstedt, M. S., et al. (2009) Breakpoint characterization of a novel approximately 59 kb genomic deletion on 19q13.42 in autosomal-dominant retinitis pigmentosa with incomplete penetrance. *Eur J Hum Genet* 17(5): 651-655.

Kozarewa, I., Armisen, J., Gardner, A. F., Slatko, B. E. & Hendrickson, C. L. (2015) Overview of Target Enrichment Strategies. *Curr Protoc Mol Biol* 112: 7 21 21-23.

Kranich, H., Bartkowski, S., Denton, M. J., Krey, S., Dickinson, P., et al. (1993) Autosomal dominant 'sector' retinitis pigmentosa due to a point mutation predicting an Asn-15-Ser substitution of rhodopsin. *Hum Mol Genet* 2(6): 813-814.

L

- Latasiewicz, M., Salvetti, A. P. & MacLaren, R. E. (2017) A novel mutation in the dominantly inherited TOPORS gene supports haploinsufficiency as the mechanism of retinitis pigmentosa. *Ophthalmic Genet* 38(6): 562-566.
- Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14): 1754-1760.
- Li, H. & Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5): 589-595.
- Liew, G., Michaelides, M. & Bunce, C. (2014) A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. *BMJ Open* 4(2): e004015.
- Lim, K. P., Yip, S. P., Cheung, S. C., Leung, K. W., Lam, S. T. & To, C. H. (2009) Novel PRPF31 and PRPH2 mutations and co-occurrence of PRPF31 and RHO mutations in Chinese patients with retinitis pigmentosa. *Arch Ophthalmol* 127(6): 784-790.
- Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., et al. (2007) Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science* 316(5821): 115-120.
- Liu, Y., Chen, X., Xu, Q., Gao, X., Tam, P. O., et al. (2015) SPP2 Mutations Cause Autosomal Dominant Retinitis Pigmentosa. *Sci Rep* 5: 14867.
- Lledo-Riquelme, M., Campos-Mollo, E. & Cuenca, N. (2010) La transducción visual. *Annals d'Oftalmologia* 18(3):130-136.

M

- Ma, K. K., Lin, J., Boudreault, K., Chen, R. W. & Tsang, S. H. (2017) PHENOTYPING CHOROIDEREMIA AND ITS CARRIER STATE WITH MULTIMODAL IMAGING TECHNIQUES. *Retin Cases Brief Rep* 11 Suppl 1: S178-S181.
- MacDonald, I. M., Hume, S., Chan, S. & Seabra, M. C. (1993) Choroideremia.
- Makarova, O. V., Makarov, E. M., Liu, S., Vornlocher, H. P. & Luhrmann, R. (2002) Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *EMBO J* 21(5): 1148-1157.
- Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., et al. (2010) Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7(2): 111-118.
- Martin-Merida, I., Sanchez-Alcudia, R., Fernandez-San Jose, P., Blanco-Kelly, F., Perez-Carro, R., et al. (2017) Analysis of the PRPF31 Gene in Spanish Autosomal Dominant Retinitis

- Pigmentosa Patients: A Novel Genomic Rearrangement. *Invest Ophthalmol Vis Sci* 58(2): 1045-1053.
- Martinez-Gimeno, M., Trujillo, M. J., Lorda, I., Gimenez, A., Calvo, M. T., et al. (2000) Three novel mutations (P215L, T289P, and 3811-2 A-->G) in the rhodopsin gene in autosomal dominant retinitis pigmentosa in Spanish families. *Hum Mutat* 16(1): 95-96.
- Martinez-Gimeno, M., Maseras, M., Baiget, M., Beneito, M., Antinolo, G., et al. (2001) Mutations P51U and G122E in retinal transcription factor NRL associated with autosomal dominant and sporadic retinitis pigmentosa. *Hum Mutat* 17(6): 520.
- Martinez-Gimeno, M., Gamundi, M. J., Hernan, I., Maseras, M., Milla, E., et al. (2003) Mutations in the pre-mRNA splicing-factor genes PRPF3, PRPF8, and PRPF31 in Spanish families with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 44(5): 2171-2177.
- Matthijs, G., Souche, E., Alders, M., Corveleyn, A., Eck, S., et al. (2016) Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet.* 24(1):2-5.
- McGee, T. L., Devoto, M., Ott, J., Berson, E. L. & Dryja, T. P. (1997) Evidence that the penetrance of mutations at the RP11 locus causing dominant retinitis pigmentosa is influenced by a gene linked to the homologous RP11 allele. *Am J Hum Genet* 61(5): 1059-1066.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9): 1297-1303.
- McKie, A. B., McHale, J. C., Keen, T. J., Tarttelin, E. E., Goliath, R., et al. (2001) Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum Mol Genet* 10(15): 1555-1562.
- McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P. & Cunningham, F. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 26(16): 2069-2070.
- Mendes, H. F., van der Spuy, J., Chapple, J. P. & Cheetham, M. E. (2005) Mechanisms of cell death in rhodopsin retinitis pigmentosa: implications for therapy. *Trends Mol Med* 11(4): 177-185.
- N
- Neidhardt, J., Glaus, E., Lorenz, B., Netzer, C., Li, Y., et al. (2008) Identification of novel mutations in X-linked retinitis pigmentosa families and implications for diagnostic testing. *Mol Vis* 14: 1081-1093.
- Nishiguchi, K. M., Tearle, R. G., Liu, Y. P., Oh, E. C., Miyake, N., et al. (2013) Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and NEK2 as a new disease gene. *Proc Natl Acad Sci U S A* 110(40): 16139-16144.

O

- Oishi, M., Oishi, A., Gotoh, N., Ogino, K., Higasa, K., et al. (2014) Comprehensive molecular diagnosis of a large cohort of Japanese retinitis pigmentosa and Usher syndrome patients by next-generation sequencing. *Invest Ophthalmol Vis Sci* 55(11): 7369-7375.
- O'Roak, B. J., Vives, L., Fu, W., Egertson, J. D., Stanaway, I. B., et al. (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338(6114): 1619-1622.

P

- Paloma, E., Martinez-Mir, A., Vilageliu, L., Gonzalez-Duarte, R. & Balcells, S. (2001) Spectrum of ABCA4 (ABCR) gene mutations in Spanish patients with autosomal recessive macular dystrophies. *Hum Mutat* 17(6): 504-510.
- Paloma, E., Martinez-Mir, A., Garcia-Sandoval, B., Ayuso, C., Vilageliu, L., et al. (2002) Novel homozygous mutation in the alpha subunit of the rod cGMP gated channel (CNGA1) in two Spanish sibs affected with autosomal recessive retinitis pigmentosa. *J Med Genet* 39(10): E66.
- Perez-Cano, H. J., Garnica-Hayashi, R. E. & Zenteno, J. C. (2009) CHM gene molecular analysis and X-chromosome inactivation pattern determination in two families with choroideremia. *Am J Med Genet A* 149A(10): 2134-2140.
- Perez-Carro, R., Corton, M., Sanchez-Navarro, I., Zurita, O., Sanchez-Bolivar, N., et al. (2016) Panel-based NGS Reveals Novel Pathogenic Mutations in Autosomal Recessive Retinitis Pigmentosa. *Sci Rep* 6: 19531.
- Pérez-Carro, R. (2017). *Implementación de las técnicas de secuenciación masiva para la elaboración de nuevos algoritmos diagnósticos e identificación de nuevos genes en distrofias de retina* (Tesis Doctoral). Universidad Autónoma de Madrid, Madrid.
- Pierce, E. A., Quinn, T., Meehan, T., McGee, T. L., Berson, E. L. & Dryja, T. P. (1999) Mutations in a gene encoding a new oxygen-regulated photoreceptor protein cause dominant retinitis pigmentosa. *Nat Genet* 22(3): 248-254.
- Pomares, E., Riera, M., Permanyer, J., Mendez, P., Castro-Navarro, J., et al. (2010) Comprehensive SNP-chip for retinitis pigmentosa-Leber congenital amaurosis diagnosis: new mutations and detection of mutational founder effects. *Eur J Hum Genet* 18(1): 118-124.

R

- Reig, C., Antich, J., Gean, E., Garcia-Sandoval, B., Ramos, C., et al. (1994) Identification of a novel rhodopsin mutation (Met-44-Thr) in a simplex case of retinitis pigmentosa. *Hum Genet* 94(3): 283-286.

- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., et al. (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 17 (5):405-24.
- Rio Frio, T., Wade, N. M., Ransijn, A., Berson, E. L., Beckmann, J. S. & Rivolta, C. (2008) Premature termination codons in PRPF31 cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. *J Clin Invest* 118(4): 1519-1531.
- Riveiro-Alvarez, R., Aguirre-Lamban, J., Lopez-Martinez, M. A., Trujillo-Tiebas, M. J., Cantalapiedra, D., et al. (2009) Frequency of ABCA4 mutations in 278 Spanish controls: an insight into the prevalence of autosomal recessive Stargardt disease. *Br J Ophthalmol* 93(10): 1359-1364.
- Rivolta, C., McGee, T. L., Rio Frio, T., Jensen, R. V., Berson, E. L. & Dryja, T. P. (2006) Variation in retinitis pigmentosa-11 (PRPF31 or RP11) gene expression between symptomatic and asymptomatic patients with dominant RP11 mutations. *Hum Mutat* 27(7): 644-653.
- Roberts, L., Ramesar, R. & Greenberg, J. (2000) Low frequency of rhodopsin mutations in South African patients with autosomal dominant retinitis pigmentosa. *Clin Genet* 58(1): 77-78.
- Rose, A. M., Mukhopadhyay, R., Webster, A. R., Bhattacharya, S. S. & Waseem, N. H. (2011) A 112 kb deletion in chromosome 19q13.42 leads to retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 52(9): 6597-6603.
- Rose, A. M., Shah, A. Z., Venturini, G., Rivolta, C., Rose, G. E. & Bhattacharya, S. S. (2014) Dominant PRPF31 mutations are hypostatic to a recessive CNOT3 polymorphism in retinitis pigmentosa: a novel phenomenon of "linked trans-acting epistasis". *Ann Hum Genet* 78(1): 62-71.
- Rose, A. M. & Bhattacharya, S. S. (2016) Variant haploinsufficiency and phenotypic non-penetrance in PRPF31-associated retinitis pigmentosa. *Clin Genet*.
- Rose, A. M., Shah, A. Z., Venturini, G., Krishna, A., Chakravarti, A., et al. (2016) Transcriptional regulation of PRPF31 gene expression by MSR1 repeat elements causes incomplete penetrance in retinitis pigmentosa. *Sci Rep* 6: 19450.
- S
- Sato, H., Wada, Y., Itabashi, T., Nakamura, M., Kawamura, M. & Tamai, M. (2005) Mutations in the pre-mRNA splicing gene, PRPF31, in Japanese families with autosomal dominant retinitis pigmentosa. *Am J Ophthalmol* 140(3): 537-540.
- Schatz, P., Preising, M., Lorenz, B., Sander, B., Larsen, M., et al. (2010) Lack of autofluorescence in fundus albipunctatus associated with mutations in RDH5. *Retina* 30(10): 1704-1713.

- Shanks, M. E., Downes, S. M., Copley, R. R., Lise, S., Broxholme, J., et al. (2013) Next-generation sequencing (NGS) as a diagnostic tool for retinal degeneration reveals a much higher detection rate in early-onset disease. *Eur J Hum Genet* 21(3): 274-280.
- Sharon, D., Bruns, G. A., McGee, T. L., Sandberg, M. A., Berson, E. L. & Dryja, T. P. (2000) X-linked retinitis pigmentosa: mutation spectrum of the RPGR and RP2 genes and correlation with visual function. *Invest Ophthalmol Vis Sci* 41(9): 2712-2721.
- Sheffield, V. C., Fishman, G. A., Beck, J. S., Kimura, A. E. & Stone, E. M. (1991) Identification of novel rhodopsin mutations associated with retinitis pigmentosa by GC-clamped denaturing gradient gel electrophoresis. *Am J Hum Genet* 49(4): 699-706.
- Sohocki, M. M., Sullivan, L. S., Mintz-Hittner, H. A., Birch, D., Heckenlively, J. R., et al. (1998) A range of clinical phenotypes associated with mutations in CRX, a photoreceptor transcription-factor gene. *Am J Hum Genet* 63(5): 1307-1315.
- Sohocki, M. M., Daiger, S. P., Bowne, S. J., Rodriguez, J. A., Northrup, H., et al. (2001) Prevalence of mutations causing retinitis pigmentosa and other inherited retinopathies. *Hum Mutat* 17(1): 42-51.
- Sorrentino, F. S., Gallenga, C. E., Bonifazzi, C. & Perri, P. (2016) A challenge to the striking genotypic heterogeneity of retinitis pigmentosa: a better understanding of the pathophysiology using the newest genetic strategies. *Eye (Lond)* 30(12): 1542-1548.
- Strom, S. P., Clark, M. J., Martinez, A., Garcia, S., Abdelazeem, A. A., et al. (2016) De Novo Occurrence of a Variant in ARL3 and Apparent Autosomal Dominant Transmission of Retinitis Pigmentosa. *PLoS One* 11(3): e0150944.
- Sullivan, L. S., Bowne, S. J., Birch, D. G., Highbanks-Wheaton, D., Heckenlively, J. R., et al. (2006a) Prevalence of disease-causing mutations in families with autosomal dominant retinitis pigmentosa: a screen of known genes in 200 families. *Invest Ophthalmol Vis Sci* 47(7): 3052-3064.
- Sullivan, L. S., Bowne, S. J., Seaman, C. R., Blanton, S. H., Lewis, R. A., et al. (2006b) Genomic rearrangements of the PRPF31 gene account for 2.5% of autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 47(10): 4579-4588.
- Sullivan, L. S., Koboldt, D. C., Bowne, S. J., Lang, S., Blanton, S. H., et al. (2014) A dominant mutation in hexokinase 1 (HK1) causes retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 55(11): 7147-7158.
- Sung, C. H., Davenport, C. M., Hennessey, J. C., Maumenee, I. H., Jacobson, S. G., et al. (1991) Rhodopsin mutations in autosomal dominant retinitis pigmentosa. *Proc Natl Acad Sci U S A* 88(15): 6481-6485.
- Swain, P. K., Chen, S., Wang, Q. L., Affatigato, L. M., Coats, C. L., et al. (1997) Mutations in the cone-rod homeobox gene are associated with the cone-rod dystrophy photoreceptor degeneration. *Neuron* 19(6): 1329-1336.

T

- Tanackovic, G., Ransijn, A., Thibault, P., Abou Elela, S., Klinck, R., et al. (2011) PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum Mol Genet* 20(11): 2116-2130.
- Tarn, W. Y. & Steitz, J. A. (1997) Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends Biochem Sci* 22(4): 132-137.
- Trujillo, M. J., Bueno, J., Osorio, A., Sanz, R., Garcia-Sandoval, B., et al. (1998) Three novel RDS-peripherin mutations (689delT, 857del17, G208D) in Spanish families affected with autosomal dominant retinal degenerations. Mutations in brief no. 147. Online. *Hum Mutat* 12(1): 70.

V

- Vaithinathan, R., Berson, E. L. & Dryja, T. P. (1994) Further screening of the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa. *Genomics* 21(2): 461-463.
- Van Cauwenbergh, C., Coppieters, F., Roels, D., De Jaegere, S., Flipts, H., et al. (2017) Mutations in Splicing Factor Genes Are a Major Cause of Autosomal Dominant Retinitis Pigmentosa in Belgian Families. *PLoS One* 12(1): e0170038.
- Van Schil, K., Naessens, S., Van de Sompele, S., Carron, M., Aslanidis, A., et al. (2017) Mapping the genomic landscape of inherited retinal disease genes prioritizes genes prone to coding and noncoding copy-number variations. *Genet Med*.
- Venturini, G., Rose, A. M., Shah, A. Z., Bhattacharya, S. S. & Rivolta, C. (2012) CNOT3 is a modifier of PRPF31 mutations in retinitis pigmentosa with incomplete penetrance. *PLoS Genet* 8(11): e1003040.
- Vervoort, R., Lennon, A., Bird, A. C., Tulloch, B., Axton, R., et al. (2000) Mutational hot spot within a new RPGR exon in X-linked retinitis pigmentosa. *Nat Genet* 25(4): 462-466.
- Vithana, E. N., Abu-Safieh, L., Allen, M. J., Carey, A., Papaioannou, M., et al. (2001) A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell* 8(2): 375-381.
- Vithana, E. N., Abu-Safieh, L., Pelosini, L., Winchester, E., Hornan, D., et al. (2003) Expression of PRPF31 mRNA in patients with autosomal dominant retinitis pigmentosa: a molecular clue for incomplete penetrance? *Invest Ophthalmol Vis Sci* 44(10): 4204-4209.

W

- Wang, D. Y., Chan, W. M., Tam, P. O., Baum, L., Lam, D. S., et al. (2005) Gene mutations in retinitis pigmentosa and their clinical implications. *Clin Chim Acta* 351(1-2): 5-16.

Wang, K., Li, M. & Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16): e164.

Waseem, N. H., Vaclavik, V., Webster, A., Jenkins, S. A., Bird, A. C. & Bhattacharya, S. S. (2007) Mutations in the gene coding for the pre-mRNA splicing factor, PRPF31, in patients with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 48(3): 1330-1334.

Webb, T. R., Parfitt, D. A., Gardner, J. C., Martinez, A., Bevilacqua, D., et al. (2012) Deep intronic mutation in OFD1, identified by targeted genomic next-generation sequencing, causes a severe form of X-linked retinitis pigmentosa (RP23). *Hum Mol Genet* 21(16): 3647-3654.

X

Xu, F., Sui, R., Liang, X., Li, H., Jiang, R. & Dong, F. (2012) Novel PRPF31 mutations associated with Chinese autosomal dominant retinitis pigmentosa patients. *Mol Vis* 18: 3021-xxx.

Xu, Y., Guan, L., Shen, T., Zhang, J., Xiao, X., et al. (2014) Mutations of 60 known causative genes in 157 families with retinitis pigmentosa based on exome sequencing. *Hum Genet* 133(10): 1255-1271.

Y

Yanagihashi, S., Nakazawa, M., Kurotaki, J., Sato, M., Miyagawa, Y. & Ohguro, H. (2003) Autosomal dominant central areolar choroidal dystrophy and a novel Arg195Leu mutation in the peripherin/RDS gene. *Arch Ophthalmol* 121(10): 1458-1461.

Yang, Z., Chen, Y., Lillo, C., Chien, J., Yu, Z., et al. (2008) Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice. *J Clin Invest* 118(8): 2908-2916.

Yang, G., Xie, S., Feng, N., Yuan, Z., Zhang, M. & Zhao, J. (2014) Spectrum of rhodopsin gene mutations in Chinese patients with retinitis pigmentosa. *Mol Vis* 20: 1132-1136.

Z

Zernant, J., Xie, Y. A., Ayuso, C., Riveiro-Alvarez, R., Lopez-Martinez, M. A., et al. (2014) Analysis of the ABCA4 genomic locus in Stargardt disease. *Hum Mol Genet* 23(25): 6797-6806.

Zhang, J., Wang, C., Shen, Y., Chen, N., Wang, L., et al. (2016) A mutation in ADIPOR1 causes nonsyndromic autosomal dominant retinitis pigmentosa. *Hum Genet* 135(12): 1375-1387.

Zhang, Q., Xu, M., Verriotto, J. D., Li, Y., Wang, H., et al. (2016) Next-generation sequencing-based molecular diagnosis of 35 Hispanic retinitis pigmentosa probands. *Sci Rep* 6: 32792.

- Zhao, C., Bellur, D. L., Lu, S., Zhao, F., Grassi, M. A., et al. (2009) Autosomal-dominant retinitis pigmentosa caused by a mutation in SNRNP200, a gene required for unwinding of U4/U6 snRNAs. *Am J Hum Genet* 85(5): 617-627.
- Ziviello, C., Simonelli, F., Testa, F., Anastasi, M., Marzoli, S. B., et al. (2005) Molecular genetics of autosomal dominant retinitis pigmentosa (ADRP): a comprehensive study of 43 Italian families. *J Med Genet* 42(7): e47.

Anexos

Anexo1. Listado de SNVs y pequeñas in/dels probablemente patogénicas identificadas en nuestra cohorte. En negrita: variantes descritas en nuestra cohorte por primera vez.

Gen	Exón	Cambio Nucleotídico	Cambio Proteína	No. Familias	Referencia
Genes con herencia autosómica dominante, fenotipo de retinosis pigmentaria					
<i>BEST1</i>	6	c.682G>A	p.Asp228Asn	2	(Davidson <i>et al.</i> , 2009)
<i>CRX</i>	3	c.122G>A	p.Arg41Gln	1	(Swain <i>et al.</i> , 1997)
<i>CRX</i>	4	c.429del	p.Pro145Leufs*42	1	Presente estudio
<i>CRX</i>	4	c.590dup	p.Ala198Glyfs*38	1	(Sohocki <i>et al.</i> , 1998)
<i>GUCA1B</i>	1	c.131G>A	p.Arg44His	1	(Fernandez-San Jose <i>et al.</i>, 2015a)
<i>HK1</i>	20	c.2551G>A	p.Glu851Lys	1	(Sullivan <i>et al.</i> , 2014)
<i>IMPDH1</i>	8	c.585T>G	p.Phe195Leu	1	Presente estudio
<i>IMPDH1</i>	9	c.809T>G	p.Leu270Arg	1	(Borras <i>et al.</i>, 2013)
<i>IMPDH1</i>	10	c.926G>C	p.Arg309Pro	2	(Kennan <i>et al.</i> , 2002)
<i>IMPDH1</i>	10	c.931G>A	p.Asp311Asn	1	(Bowne <i>et al.</i> , 2002)
<i>IMPDH1</i>	10	c.962C>T	p.Ala321Val	1	(Fernandez-San Jose <i>et al.</i>, 2015a)
<i>NR2E3</i>	2	c.166G>A	p.Gly56Arg	7	(Coppieters <i>et al.</i> , 2007)
<i>NRL</i>	2	c.152C>T	p.Pro51Leu	2	(Martinez-Gimeno <i>et al.</i>, 2001)
<i>NRL</i>	2	c.287T>C	p.Met96Thr	1	(Hernan <i>et al.</i> , 2012)
<i>PROM1</i>	12	c.1117C>T	p.Arg373Cys	1	(Z. Yang <i>et al.</i> , 2008)
<i>PRPF3</i>	11	c.1481C>T	p.Thr494Met	3	(Chakarova <i>et al.</i> , 2002)
<i>PRPF31</i>	2	c.55del	p.Glu19Lysfs*46	1	Presente estudio
<i>PRPF31</i>	4	c.322+4_322+7del	<i>Splicing</i>	2	(Q. Zhang <i>et al.</i> , 2016)
<i>PRPF31</i>	5	c.328_330del	p.Ile110del	1	(de Sousa Dias <i>et al.</i>, 2013)
<i>PRPF31</i>	IVS6+1	c.527+1G>T	<i>Splicing</i>	1	(Chakarova <i>et al.</i> , 2006)
<i>PRPF31</i>	7	c.541G>T	p.Glu181*	2	(Pomares <i>et al.</i> , 2010)
<i>PRPF31</i>	8	c.736G>A	p.Ala246Thr	1	(Xu <i>et al.</i> , 2014)
<i>PRPF31</i>	8	c.770dup	p.Thr258Aspfs*21	1	(Vithana <i>et al.</i> , 2001)
<i>PRPF31</i>	8	c.828_829del	p.His276Glnfs*2	1	(Martinez-Gimeno <i>et al.</i>, 2003)
<i>PRPF31</i>	9	c.895T>C	p.Cys299Arg	1	(Sullivan <i>et al.</i> , 2006a)
<i>PRPF31</i>	9	c.939dup	p.Gly314Argfs*10	1	(Fernandez-San Jose <i>et al.</i>, 2015a)
<i>PRPF31</i>	IVS10-1	c.1074-1G>T	<i>Splicing</i>	1	(Martin-Merida <i>et al.</i>, 2017)
<i>PRPF31</i>	IVS11+2	c.1146+2T>C	<i>Splicing</i>	1	(Waseem <i>et al.</i> , 2007)
<i>PRPF31</i>	IVS11+2	c.1146+2T>A	<i>Splicing</i>	1	(Martin-Merida <i>et al.</i>, 2017)
<i>PRPF31</i>	14	c.1462_1472del	p.Lys488Argfs*75	1	Presente estudio
<i>PRPF6</i>	IVS17+1	c.2339+1G>A	<i>Splicing</i>	1	Presente estudio
<i>PRPF8</i>	43	c.6893_6896delinsCCATAGA	p.Leu2298_Ala2299delinsProlleGlu	1	(Martinez-Gimeno <i>et al.</i>, 2003)
<i>PRPF8</i>	43	c.6974_6994del	p.Val2325_Glu2331del	1	(Martinez-Gimeno <i>et al.</i>, 2003)
<i>PRPF8</i>	43	c.6928A>G	p.Arg2310Gly	1	(McKie <i>et al.</i> , 2001)
<i>PRPH2</i>	1	c.536G>T	p.Trp179Leu	1	(Kitiratschky <i>et al.</i> , 2011)
<i>PRPH2</i>	1	c.556G>A	p.Asp186Asn	1	(Fernandez-San Jose <i>et al.</i>, 2015a)
<i>PRPH2</i>	IVS1-1	c.582-1G>A	<i>Splicing</i>	2	(Fernandez-San Jose <i>et al.</i>, 2015a)
<i>PRPH2</i>	2	c.584G>T	p.Arg195Leu	1	(Yanagihashi <i>et al.</i> , 2003)
<i>PRPH2</i>	2	c.623G>A	p.Gly208Asp	1	(Kohl <i>et al.</i> , 1997)
<i>PRPH2</i>	2	c.634A>G	p.Ser212Gly	2	(Farrar <i>et al.</i> , 1992)
<i>PRPH2</i>	2	c.643A>T	p.Asn215Tyr	1	Presente estudio
<i>PRPH2</i>	2	c.646C>T	p.Pro216Ser	1	(Fishman <i>et al.</i> , 1994)

Anexo1. Espectro de SNVs probablemente patogénicas identificadas en nuestra cohorte (continuación).

Gen	Exón	Cambio Nucleotídico	Cambio Proteína	No. Familias	Referencia
<i>RHO</i>	1	c.44A>G	p.Asn15Ser	1	(Kranich <i>et al.</i> , 1993)
<i>RHO</i>	1	c.50C>T	p.Thr17Met	2	(Sheffield <i>et al.</i> , 1991)
<i>RHO</i>	1	c.84G>C	p.Gln28His	2	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	1	c.131T>C	p.Met44Thr	1	(Reig <i>et al.</i> , 1994)
<i>RHO</i>	1	c.173C>G	p.Thr58Arg	2	(Dryja <i>et al.</i> , 1990)
<i>RHO</i>	1	c.180C>A	p.Tyr60*	1	(Eisenberger <i>et al.</i> , 2013)
<i>RHO</i>	1	c.236T>C	p.Leu79Pro	1	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	1	c.290C>T	p.Thr97Ile	1	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	1	c.316G>A	p.Gly106Arg	2	(Inglehearn <i>et al.</i> , 1992)
<i>RHO</i>	2	c.378G>A	p.Tyr126*	1	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	2	c.404G>T	p.Arg135Leu	2	(Andreasson <i>et al.</i> , 1992)
<i>RHO</i>	2	c.403C>T	p.Arg135Trp	5	(Sung <i>et al.</i> , 1991)
<i>RHO</i>	2	c.491C>A	p.Ala164Glu	3	(Vaithinathan <i>et al.</i> , 1994)
<i>RHO</i>	2	c.488T>C	p.Met163Thr	1	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	2	c.500G>A	p.Cys167Tyr	2	(Fernandez-San Jose <i>et al.</i> , 2015b)
<i>RHO</i>	2	c.509C>G	p.Pro170Arg	1	(Sohocki <i>et al.</i> , 2001)
<i>RHO</i>	2	c.512C>T	p.Pro171Leu	1	(Dryja <i>et al.</i> , 1991)
<i>RHO</i>	IVS2-2	c.531-2A>G	Splicing	1	(Martinez-Gimeno <i>et al.</i> , 2000)
<i>RHO</i>	3	c.533A>G	p.Tyr178Cys	1	(Sung <i>et al.</i> , 1991)
<i>RHO</i>	3	c.541G>A	p.Glu181Lys	2	(Dryja <i>et al.</i> , 1991)
<i>RHO</i>	3	c.544G>A	p.Gly182Ser	3	(Sheffield <i>et al.</i> , 1991)
<i>RHO</i>	3	c.545G>T	p.Gly182Val	1	(Yang <i>et al.</i> , 2014)
<i>RHO</i>	3	c.556T>C	p.Ser186Pro	1	(Dryja <i>et al.</i> , 1991)
<i>RHO</i>	3	c.562G>A	p.Gly188Arg	1	(Dryja <i>et al.</i> , 1991)
<i>RHO</i>	3	c.568G>T	p.Asp190Tyr	2	(al-Maghtheh <i>et al.</i> , 1993)
<i>RHO</i>	3	c.644C>T	p.Pro215Leu	1	(Martinez-Gimeno <i>et al.</i> , 2000)
<i>RHO</i>	4	c.865A>C	p.Thr289Pro	1	(Martinez-Gimeno <i>et al.</i> , 2000)
<i>RHO</i>	5	c.1040C>T	p.Pro347Leu	10	(Dryja <i>et al.</i> , 1990)
<i>RP1</i>	4	c.1981G>T	p.Glu661*	1	(Fernandez-San Jose <i>et al.</i> , 2015a)
<i>RP1</i>	4	c.2029C>T	p.Arg677*	1	(Pierce <i>et al.</i> , 1999)
<i>RP1</i>	4	c.2056 C>T	p.Gln686*	1	(Gamundi <i>et al.</i> , 2006)
<i>RP1</i>	4	c.2080G>T	p.Gly694*	1	Presente estudio
<i>RP1</i>	4	c.2115delA	p.Gly706Valfs*7	1	(Gamundi <i>et al.</i> , 2006)
<i>RP1</i>	4	c.2164_2165delinsG	p.Lys722Glu*16	1	Presente estudio
<i>RP1</i>	4	c.2232T>A	p.Cys744*	1	(Bowne <i>et al.</i> , 1999)
<i>RP1</i>	4	c.2288del	p.Asn763Ilefs*12	1	(Fernandez-San Jose <i>et al.</i> , 2015a)
<i>RP1</i>	4	c.2289_2299del	p.Asn763Lysfs*14	1	Presente estudio
<i>RP1</i>	4	c.2745_2749del	p.Tyr915*	1	(Fernandez-San Jose <i>et al.</i> , 2015a)
<i>RP1</i>	4	c.4328G>A	p.Arg1443Gln	1	(Fernandez-San Jose <i>et al.</i> , 2015a)
<i>SNRNP200</i>	16	c.2041C>T	p.Arg681Cys	2	(Benaglio <i>et al.</i> , 2011)
<i>SNRNP200</i>	16	c.2042G>A	p.Arg681His	1	(Benaglio <i>et al.</i> , 2011)
<i>SNRNP200</i>	18	c.2359G>A	p.Ala787Thr	1	(Xu <i>et al.</i> , 2014)
<i>SNRNP200</i>	25	c.3260C>T	p.Ser1087Leu	2	(Zhao <i>et al.</i> , 2009)
<i>TOPORS</i>	3	c.2518_2519del	p.Ser840*	2	Presente estudio
<i>TOPORS</i>	3	c.2524dup	p.Thr842Asnfs*31	1	Presente estudio

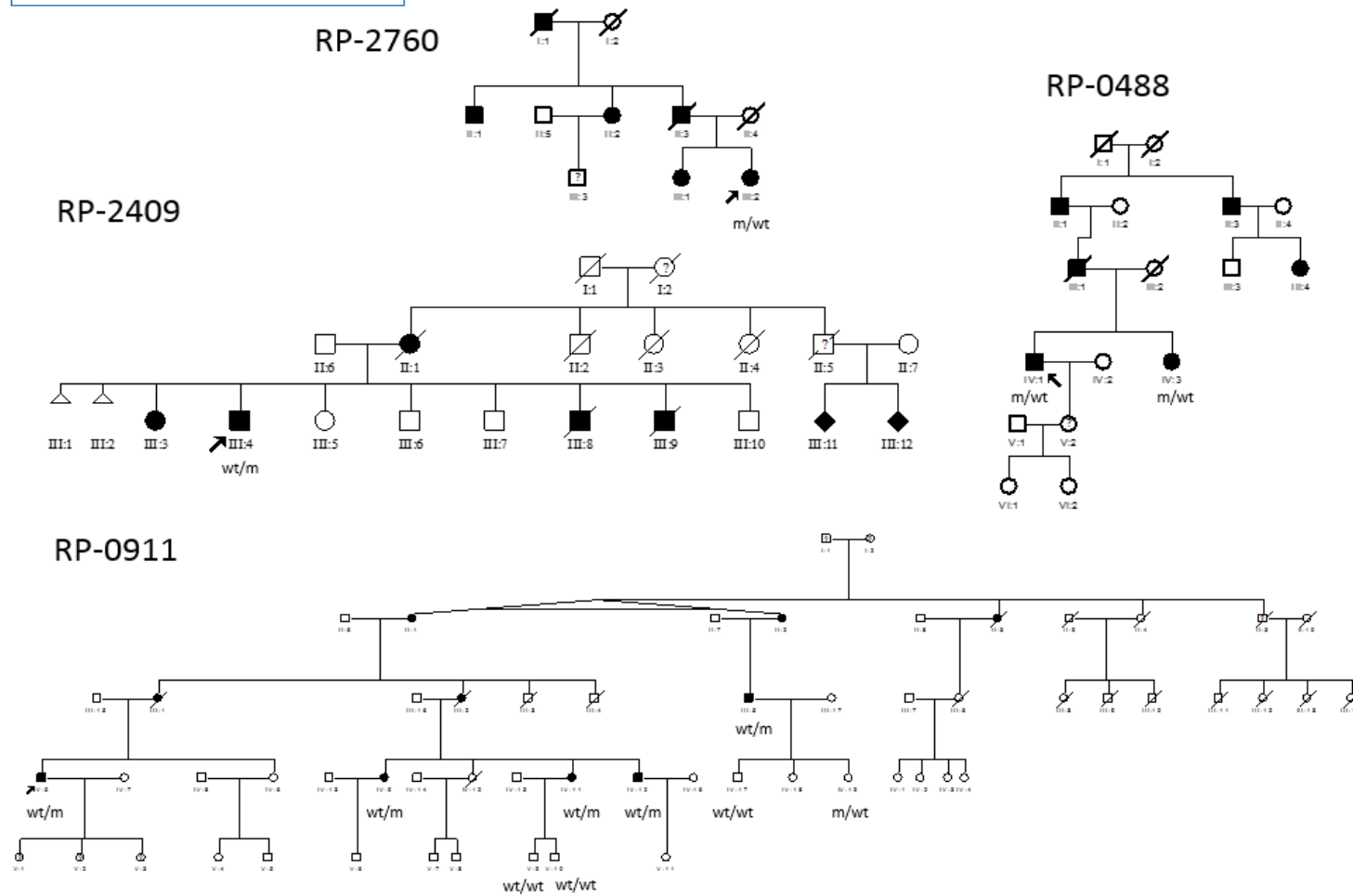
Anexo1. Espectro de SNVs probablemente patogénicas identificadas en nuestra cohorte (continuación).

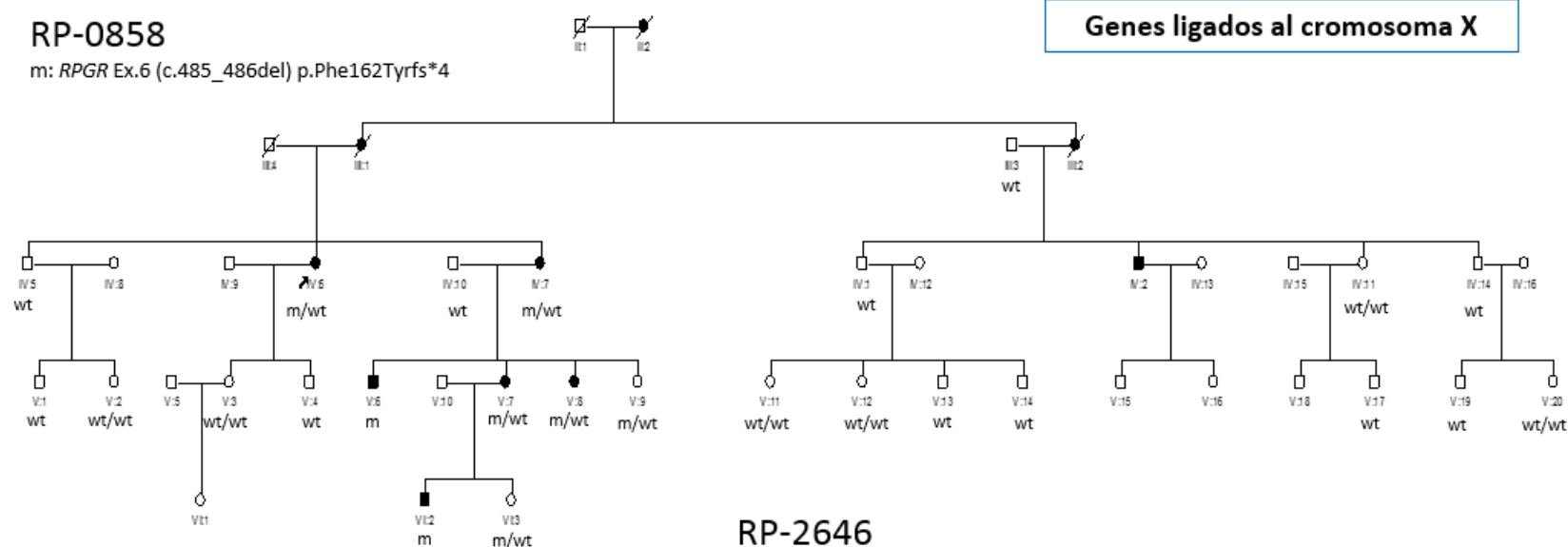
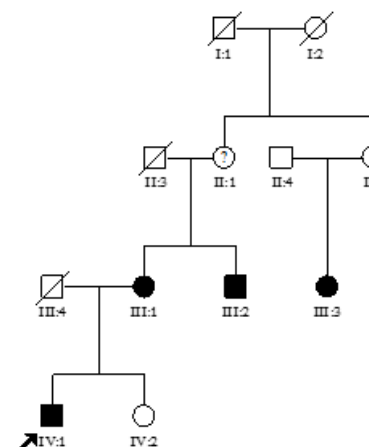
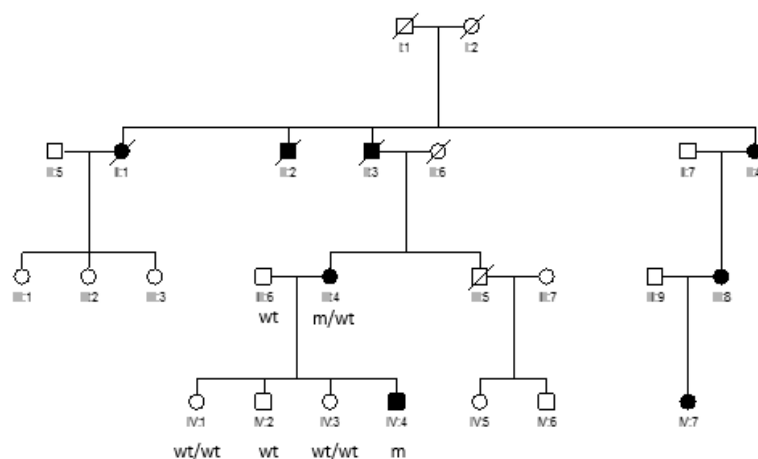
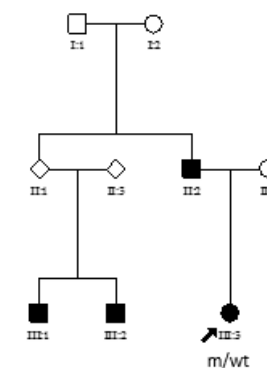
Gen	Exón	Cambio Nucleotídico	Cambio Proteína	No. Familias	Referencia
Gen con herencia autosómico dominante, fenotipo LORD (<i>late-onset retinal degeneration</i>)					
<i>C1QTNF5</i>	15	c.489C>A	p.Ser163Arg	4	(Hayward <i>et al.</i> , 2003)
Genes ligados al cromosoma X					
<i>CHM</i>	5	c.340G>T	p.Glu114*	1	Presente estudio
<i>CHM</i>	6	c.757C>T	p.Arg253*	1	(Hayakawa <i>et al.</i> , 1999)
<i>RP2</i>	1	c.1A>G	p.Met1?	1	Presente estudio
<i>RP2</i>	1	c.14_16del	p.Phe5del	1	(Neidhardt <i>et al.</i> , 2008)
<i>RPGR</i>	6	c.485_486del	p.Phe162Tyrfs*4	1	(Sharon <i>et al.</i> , 2000)
<i>RPGR</i>	8	c.888_889del	p.Ile297Lysfs*48	1	Presente estudio
<i>RPGR</i>	15	c.2405_2406del	p.Glu802Glyfs*32	1	(Vervoort <i>et al.</i> , 2000)
<i>RPGR</i>	15	c.2296_2299del	p.Gly766Asnfs*48	1	Presente estudio
Genes recesivos					
<i>ABCA4</i>	8	c.950del	p.Gly317Alafs*57	1	(Corton <i>et al.</i> , 2013)
	19	c.2888del	p.Gly963Alafs*14		(Paloma <i>et al.</i> , 2001)
	48	c.6688del	p.Leu2230Serfs*17		Presente estudio
<i>CNGA1</i>	5	c.94C>T	p.Arg32*	1	(Paloma <i>et al.</i> , 2002)
	6	c.131del	p.Glu44Glyfs*49		Presente estudio
<i>EYS</i>	IVS28	c.5928-2A>G	<i>Splicing</i>	1	(Gonzalez-del Pozo <i>et al.</i> , 2011)
<i>RDH5</i>	4	c.625C>T	p.Arg209*	1	(Schatz <i>et al.</i> , 2010)
	5	c.776C>T	p.Pro259Leu		Presente estudio
<i>USH2A</i>	13	c.2299del	p.Glu767Serfs*21	1	(Eudy <i>et al.</i> , 1998)
	62	c.12094G>A	p.Gly4032Arg		(Ge <i>et al.</i> , 2015)

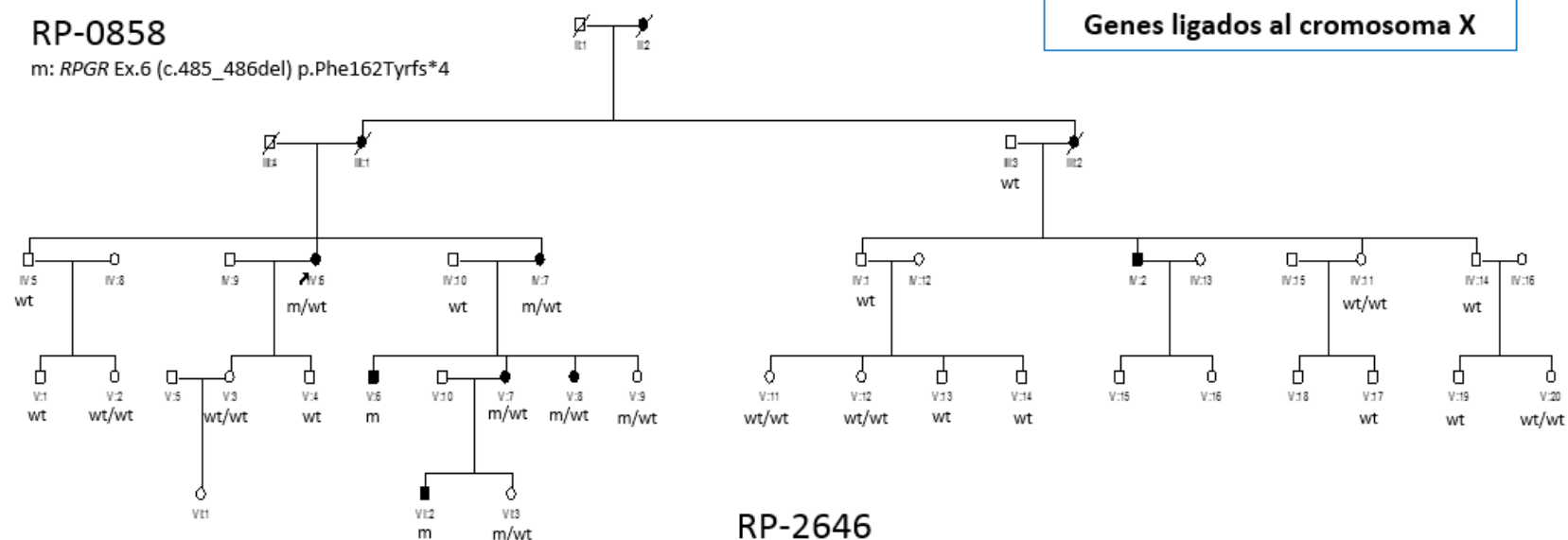
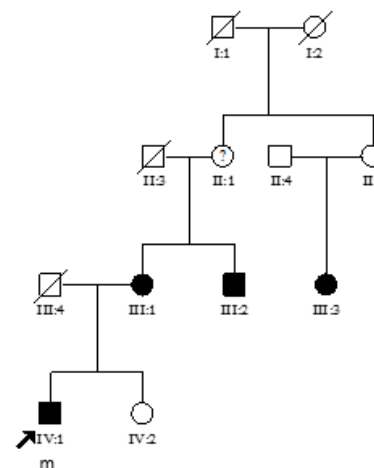
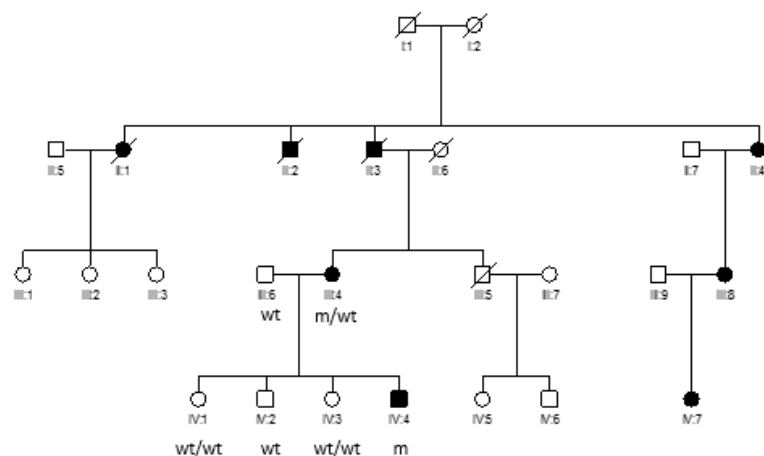
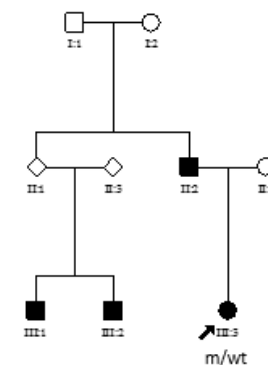
Anexo 2. Árboles y estudio de segregación de las familias reclasificadas, en el presente estudio. m: alelo mutado; wt: alelo *wild-type*.

LORD

m: C1QTNF5 Ex.15 (c.489C>A) p.Ser163Arg



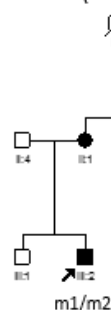
RP-0858m: *RPGR* Ex.6 (c.485_486del) p.Phe162Tyrfs*4**RP-2646**m: *RPGR* Ex.15 (c.2296_2299del) p.Gly766Asnfs*48**RP-0615**m: *RPGR* Ex.15 (c.2405_2406del) p.Glu802Glyfs*32**RP-2683**m: *RPGR* Ex.8 (c.888_889delAA) p.Ile297Lysfs*48

RP-0858m: *RPGR* Ex.6 (c.485_486del) p.Phe162Tyrfs*4**RP-2646**m: *RPGR* Ex.15 (c.2296_2299del) p.Gly766Asnfs*48**RP-0615**m: *RPGR* Ex.15 (c.2405_2406del) p.Glu802Glyfs*32**RP-2683**m: *RPGR* Ex.8 (c.888_889delAA) p.Ile297Lysfs*48

Genes con herencia recesiva

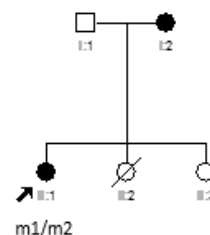
RP-1123

m1: *RDH5* Ex.4 (c.625C>T) p.Arg209*
 m2: *RDH5* Ex.5 (c.776C>T) p.Pro259Leu



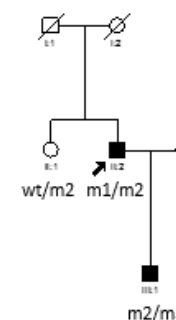
RP-1446

m1: *USH2A* Ex.13 (c.2299del) p.Glu767Serfs*21
 m2: *USH2A* Ex.62 (c.12094G>A) p.Gly4032Arg



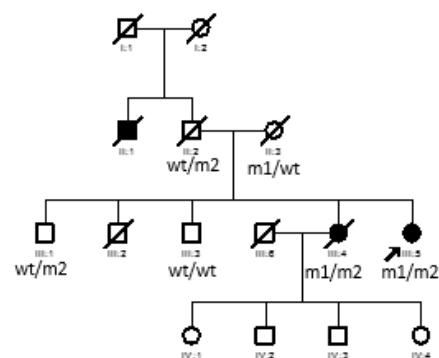
RP-1455

m1: *ABCA4* Ex.19 (c.2888delG) p.Gly963Alafs*14
 m2: *ABCA4* Ex.48 (c.6688delC) p.Leu2230Serfs*17
 m3: *ABCA4* Ex.8 (c.950delG) p.Gly317Alafs*57



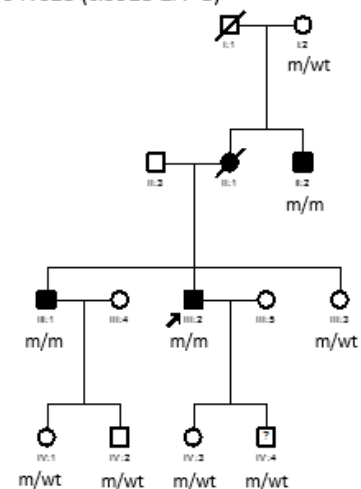
RP-0038

m1: *CNGA1* Ex.5 (c.94C>T) p.Arg32*
 m2: *CNGA1* Ex.6 (c.131delA) p.Glu44Glyfs*49

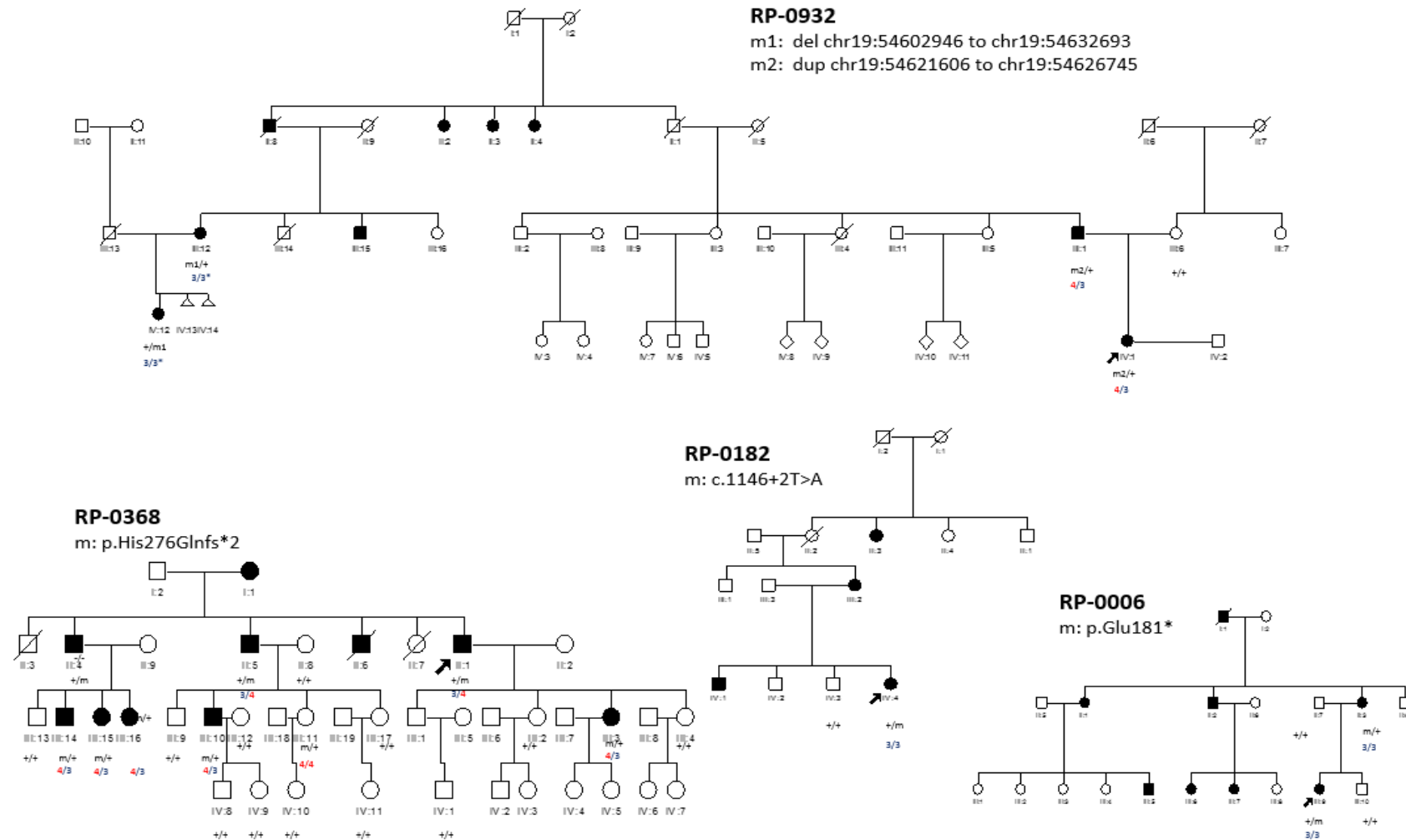


RP-1217

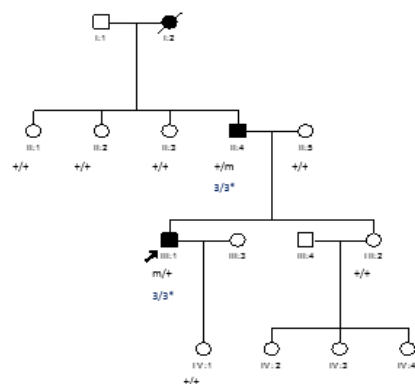
m: *EYS* IVS28 (c.5928-2A>G)



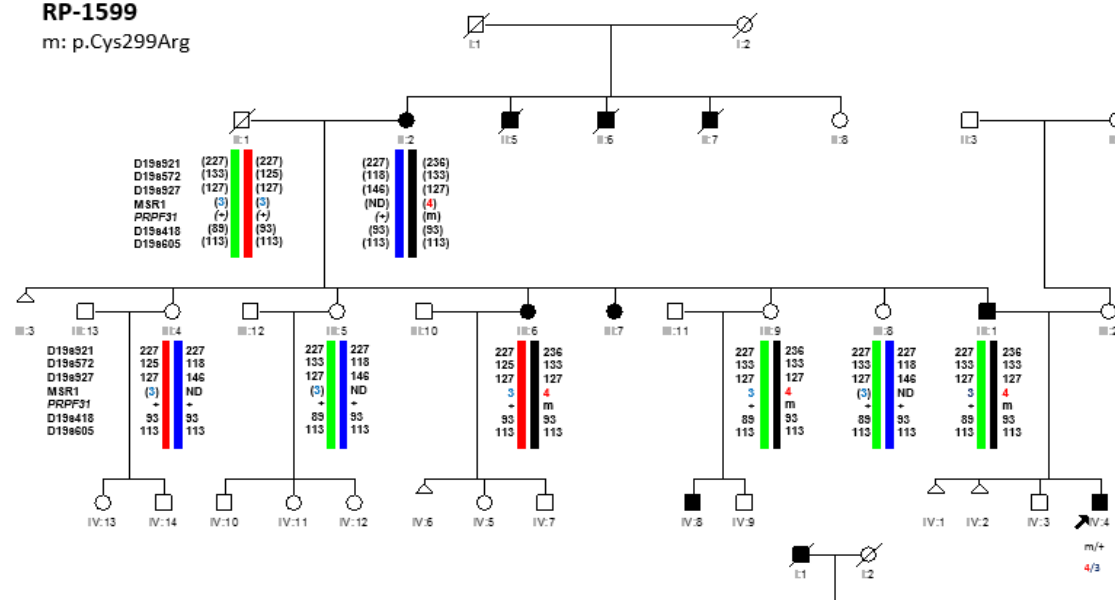
Anexo 3. Árboles y estudio de segregación de las familias adRP con mutación en el gen *PRPF31*. Resultados del estudio de CNV del minisatélite MSR1. Resultado de los haplotipos en las familias RP-0361 y RP-1599, donde el alelo de 4 copias de MSR1 se encuentra en *cis* con el alelo mutante en los portadores asintomáticos III:6 (RP-0361) y III:9 (RP-1599). m: alelo mutado; +: alelo *wild-type*.



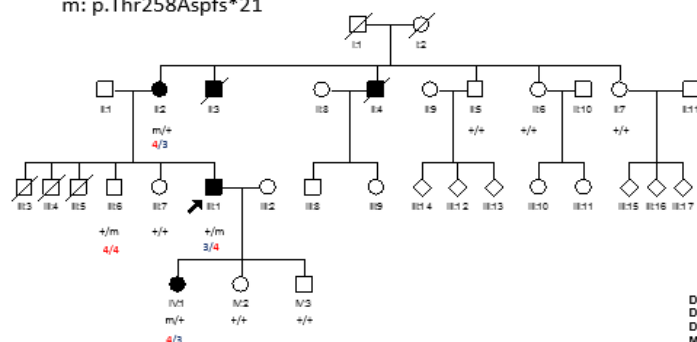
m: del chr19:54602946 to chr19:54635178



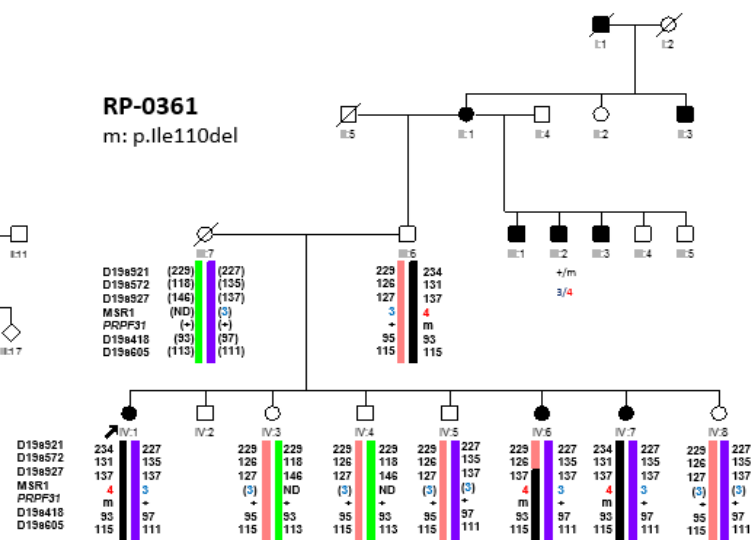
m: p.Cys299Arg



m: p.Thr258Aspfs*21

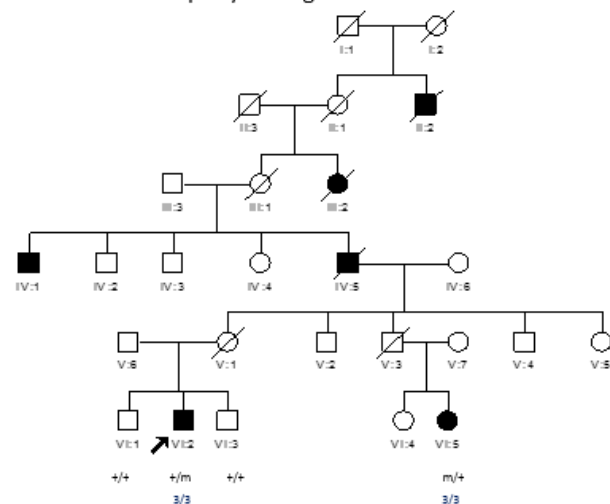


m: p.Ile110del



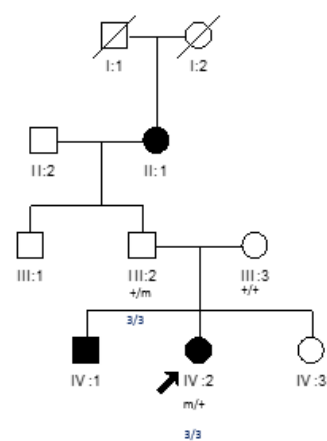
RP-1541

m: p.Gly314Argfs*10



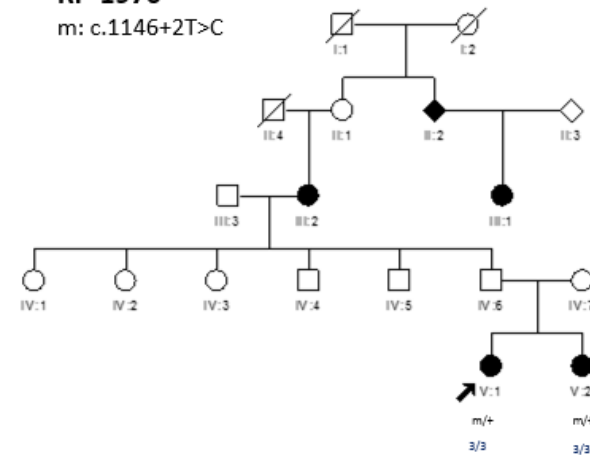
RP-1609

m: p.Glu181*



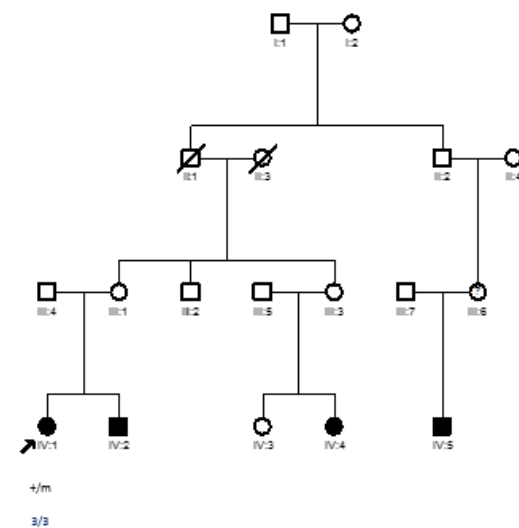
RP-1970

m: c.1146+2T>C



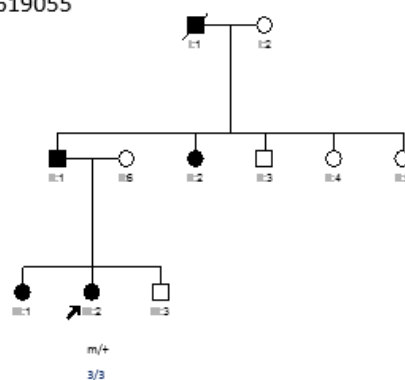
RP-1318

m: c.322+4_322+7del



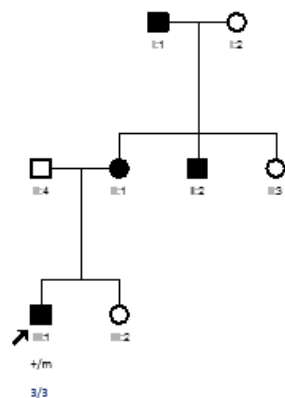
RP-2496

m: c.1074-1G>T



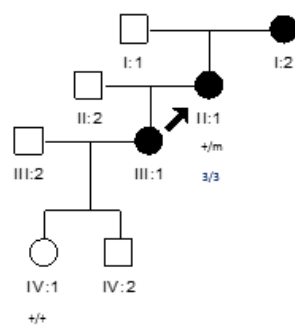
RP-2426

del chr19:54618828 to chr19:54619055



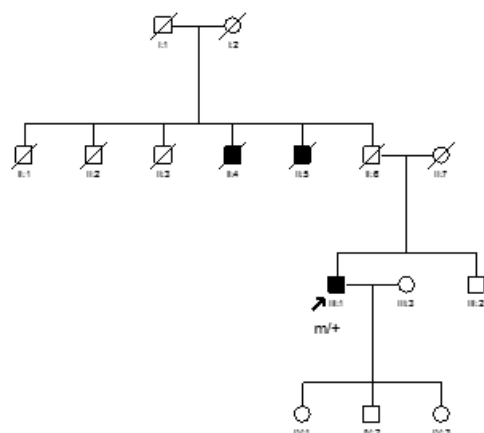
RP-2136

m: c.527+1G>T

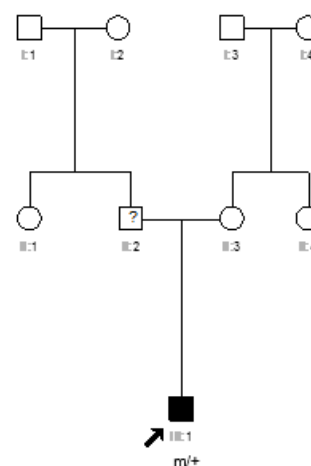


RP-2266

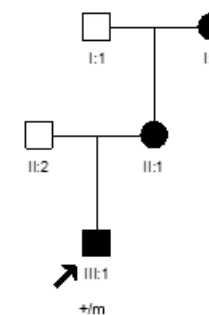
m: p.Lys488Argfs*75

**RP-2595**

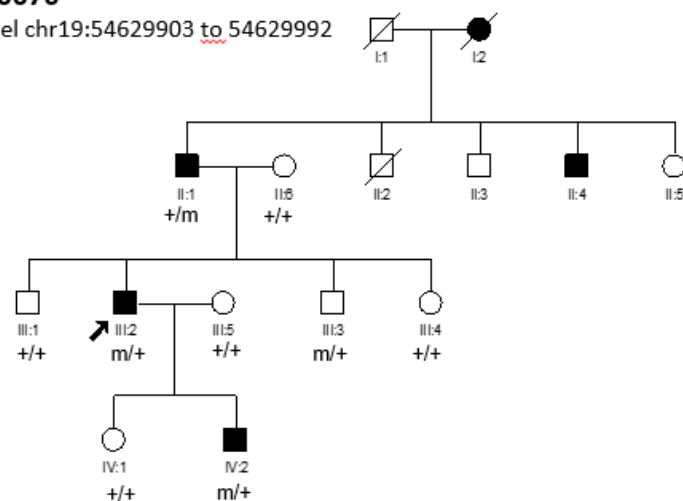
m: p.Ala246Thr

**RP-2712**

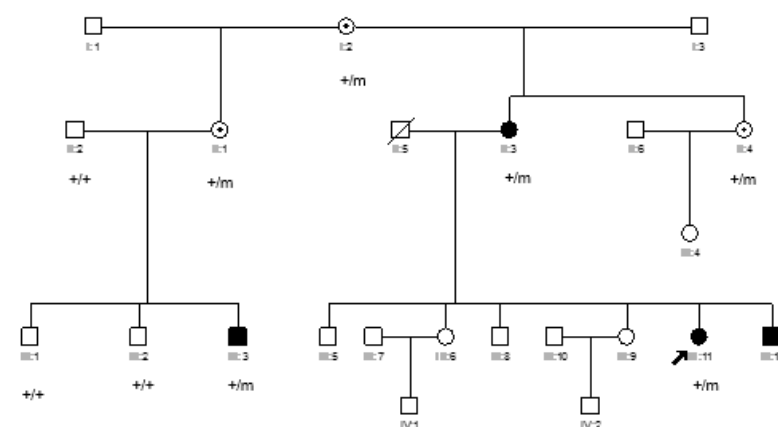
m: p.Glu19Lysfs*46

**RP-0076**

m: del chr19:54629903 to 54629992

**RP-0187**

m: c.322+4_322+7del



Anexo 4. A. Publicación derivada de esta tesis doctoral.

Genetics

Analysis of the *PRPF31* Gene in Spanish Autosomal Dominant Retinitis Pigmentosa Patients: A Novel Genomic Rearrangement

Inmaculada Martin-Merida,^{1,2} Rocio Sanchez-Alcudia,^{1,2} Patricia Fernandez-San Jose,^{1,2} Fiona Blanco-Kelly,^{1,2} Raquel Perez-Carro,^{1,2} Luciana Rodriguez-Jacy da Silva,³ Berta Almoguera,⁴ Blanca Garcia-Sandoval,⁵ Maria Isabel Lopez-Molina,⁵ Almudena Avila-Fernandez,^{1,2} Miguel Carballo,⁶ Marta Corton,^{1,2} and Carmen Ayuso^{1,2}

¹Department of Genetics, Instituto de Investigacion Sanitaria-Fundacion Jimenez Diaz University Hospital-Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

²Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain

³Universidade de Mogi das Cruzes, São Paulo, Brazil

⁴Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States

⁵Department of Ophthalmology, Instituto de Investigacion Sanitaria-Fundacion Jimenez Diaz University Hospital-Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

⁶Molecular Genetics Unit, Hospital de Terrassa, Terrassa, Barcelona, Spain

Correspondence: Carmen Ayuso, Department of Genetics, Instituto de Investigacion Sanitaria-Fundacion Jimenez Diaz University Hospital-Universidad Autónoma de Madrid (IIS-FJD, UAM), Avenue Reyes Católicos No. 2, Madrid 28040, Spain; cayuso@fjd.es.

MC and CA contributed equally to the work presented here and should therefore be regarded as equivalent authors.

Submitted: August 12, 2016

Accepted: January 15, 2017

Citation: Martin-Merida I, Sanchez-Alcudia R, Fernandez-San Jose P, et al. Analysis of the *PRPF31* gene in Spanish autosomal dominant retinitis pigmentosa patients: a novel genomic rearrangement. *Invest Ophthalmol Vis Sci.* 2017;58:1045-1053. DOI: 10.1167/jovs.16-20515

PURPOSE. The aim was to determine the prevalence of *PRPF31* mutations in a cohort of Spanish autosomal dominant retinitis pigmentosa (adRP) families to deepen knowledge of the pathogenic mechanisms underlying the disease and to assess genotype-phenotype correlations.

METHODS. A cohort of 211 adRP patients was screened for variants in *PRPF31* by using a combined strategy comprising next-generation sequencing approaches and copy-number variation (CNV) analysis. Quantitative RT-PCR and CNV analysis of the regulatory MSR1 element were also performed to assess *PRPF31* gene expression. Phenotype was assessed by using ophthalmologic examination protocols.

RESULTS. Fifteen different causative mutations and genomic rearrangements were identified, revealing five novel mutations. Prevalence of *PRPF31* mutations, genomic rearrangements, and lack of penetrance were 7.6%, 1.9%, and 66.7%, respectively. Interestingly, we identified a tandem duplication and a partial *PRPF31* deletion in different affected individuals from the same family. *PRPF31* gene expression was significantly decreased in symptomatic cases carrying either *PRPF31* duplication or deletion as compared to controls. The 4 MSR1 allele in *cis* with the *PRPF31* wild-type allele was apparently a protective factor. The mutated phenotype varied from no symptoms to typical retinitis pigmentosa with variable onset and course depending on the kind of mutation, with the duplication case the most severe.

CONCLUSIONS. In view of the high genetic heterogeneity of *PRPF31* mutations, the screening must include the entire gene, as well as CNV assays, to detect large rearrangements. This is the first report of a variable phenotype correlation as well as a gross duplication and deletion within the same family.

Keywords: *PRPF31*, adRP, CNV, NGS, MSR1

Inherited retinal dystrophies (RDs) are a group of rare diseases, most of which are degenerative and progressive as a result of photoreceptor cell death. The most common form of RD is retinitis pigmentosa (RP), with a prevalence of 1/4000.¹ Retinitis pigmentosa is characterized by rod photoreceptor degeneration causing night blindness, progressive reduction of visual field, and pigment accumulation in the peripheral retina detected on fundus examination at early stages of the disease.^{1,2} Retinitis pigmentosa shows great clinical and genetic heterogeneity, presenting autosomal recessive (35%–50%), autosomal dominant (15%–25%), or X-linked (7%–15%) inheritance patterns, making molecular diagnosis very complex.² Approx-

mately 15% of Spanish families with RP present autosomal dominant RP (adRP).³ To date, 28 genes have been associated with adRP (RetNet; available in the public domain at <https://sph.uth.tmc.edu/RetNet/>).

One of the disease-causing genes in adRP is *PRPF31* (pre-mRNA processing factor 31; NM_015629), located at 19q13.42. *PRPF31* encompasses 14 exons spanning approximately 18 kb of genomic DNA, and encoding a protein of 499 amino acids. The *PRPF31* gene is ubiquitously expressed including in neuronal tissues, brain, and retina.⁴ *PRPF31* is highly conserved from mammals to yeast⁵ and contains a snoRNA-binding domain (NOP domain). This domain is involved in the

Analysis of the *PRPF31* Gene: A Novel Genomic Rearrangement

formation of the U4/U6/U5 tri-small nuclear ribonucleoprotein, the spliceosomal key molecule.⁶ *PRPF31* plays an important role in the excision of introns from RNA transcripts during the splicing process; thus, the associated pathogenic mechanism is clearly due to the disruption of the splicing function.⁵ Mutations in this gene have only been associated with adRP. Since the retina contains the highest volume of processed pre-mRNA in the entire body, high demands on functional splicing components are required in retinal photoreceptors.⁷ Accordingly, in *PRPF31*-associated RP, haploinsufficiency of this splicing factor leads to photoreceptor death.⁷

The prevalence of *PRPF31* mutations in adRP cohorts ranges from 1% to 8% depending on geographic origin,⁸ with a previously reported prevalence of 1.7% in the Spanish population.⁹ To date, more than 100 mutations have been identified, including missense and loss-of-function (LOF) mutations, such as nonsense, splice site, and frameshift variants, and also complex genomic rearrangements, as described in Human Gene Mutation Database.¹⁰

Incomplete penetrance in mutated *PRPF31* families including asymptomatic carriers has been previously reported by McGee et al.,¹¹ presumably explained by compensation of the high expressivity form of the wild-type allele. The expression of the *PRPF31* gene has a continuous distribution in the general population with a threshold for normal retina function, so asymptomatic *PRPF31* mutation carriers usually have levels of *PRPF31* gene expression above this threshold.¹² It has been postulated that this phenomenon occurs by transcriptional regulation of different modifier genes, as *CNOT1*,^{13,14} or recently, by a minisatellite repeat element (MSR1)¹⁵ located in the *PRPF31* promoter.

The purpose of this study was to determine the prevalence of mutations in the *PRPF31* gene in a Spanish cohort of adRP families to gain knowledge of new pathogenic mechanisms underlying the disease and to assess further genotype/phenotype correlations.

SUBJECTS AND METHODS

Subjects and Samples

A total of 211 unrelated Spanish patients with initial clinical diagnosis of adRP were included in this study. These subjects were identified after a search of the Fundación Jiménez Díaz (FJD) University Hospital database, containing 25 years of collected data. Autosomal dominant RP diagnosis was based on complete ophthalmologic examination and genetic classification, as previously described.^{3,16}

All subjects provided written informed consent after the nature of the study procedures had been fully explained. The samples belong to the Biobank of the FJD University Hospital. This study was reviewed and approved by the Ethics Committee of our hospital and was performed according to the tenets of the Declaration of Helsinki and subsequent reviews.

Genomic DNA was extracted from peripheral blood leukocytes by using automated DNA extractors: BioRobot EZ1 (QIAGEN, Hilden, Germany) or MagNA Pure Compact system (Roche Applied Science, Penzberg, Germany). Total RNA was isolated from human whole blood with the PAXgene Blood RNA Kit using the automated extractor QIAcube (QIAGEN) following the manufacturer's instructions.

Mutational Analysis of *PRPF31*

Screening of the *PRPF31* gene was performed with a combined strategy of several genetic tools (Fig. 1), depending on the

IOVS | February 2017 | Vol. 58 | No. 2 | 1046

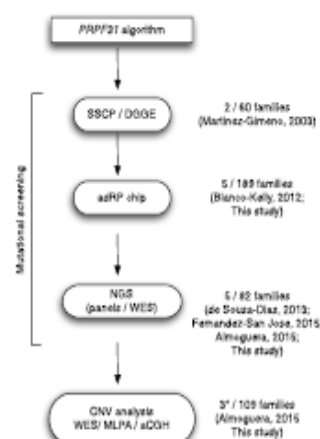


Figure 1. Molecular strategy followed for the screening of *PRPF31* in our cohort of Spanish patients with autosomal dominant retinitis pigmentosa. The workflow shows the number of families characterized versus number of families screened with each technology. *One of the families with two different rearrangements.

development of different molecular techniques. First, a screening of mutations was performed in 60 of these 211 patients by using single-strand conformation polymorphism (SSCP) and CG-clamped denaturing gradient gel electrophoresis (DGGE), as previously reported by Martínez-Gimeno et al.⁹ In addition, in 189 patients, a commercial adRP genotyping microarray (Asper Biotech, Tartu, Estonia) was applied to analyze known mutations in genes previously associated with the disease, as described by Blanco-Kelly et al.¹⁷ Second, resequencing of causative adRP genes was carried out by targeted next-generation sequencing (NGS) in a total of 76 families as previously described,^{18,19} and whole exome sequencing (WES) was performed in 6 families, as reported by Almoguer et al.²⁰ All identified variants were confirmed by Sanger sequencing and further segregated in the respective families when other relatives were available (Supplementary Fig. S1).

In silico pathogenic prediction of new splicing variants was performed with several splicing tools (Human Splicing Finder, MaxEntScan, Splice Site Finder-like, NNSPLICE, GeneSplicer, ESEFinder), all of them in the commercial Alamut software (Interactive Biosoftware, Rouen, France).

Haplotype Analysis

Five microsatellite markers with high heterozygosity (D19S921, D19S572, D19S927, D19S418, and D19S605) flanking 1.9 Mb around *PRPF31* were screened in four families. Six intragenic single-nucleotide polymorphisms (SNPs) (rs57960425, rs2303557, rs1058572, rs11556769, rs76781318, and rs655240) were genotyped in two families with the same causative mutation. Markers were separately amplified and PCR products were separated by capillary electrophoresis on an ABI 3130xl (Applied Biosystems, Foster City, CA, USA) and analyzed with GeneMapper v3.5 software (Applied Biosystems). Single-nucleotide polymorphism genotyping was performed by Sanger sequencing. Cynilic program (Cynilic Software, Wallingford, UK) was used for haplotype reconstruction.

Copy-Number Variation (CNV) Analysis

Copy-number variation analysis of *PRPF31* was performed by means of multiplex ligation-dependent probe amplification

(MLPA) analysis in 103 families with negative findings in these previous mutational screening, and genomic rearrangements were further refined by array-based comparative genomic hybridization (aCGH).

MLPA was performed according to manufacturer instructions by using the commercially available P235 Retinitis Pigmentosa Kit (MRC-Holland, Amsterdam, The Netherlands), which contains specific probes for all exons of the *PRPF31* gene. Amplified fragments were separated by capillary electrophoresis through an ABI 3130xl automatic analyzer and analyzed with GeneMapper (Applied Biosystems) and Coffalyser (MRC-Holland) software programs. The data were interpreted by taking the ratio of each probe signal between the control and the patient DNA sample. A ratio of 1.0 indicates the presence of two alleles (normal diploid), while a ratio below 0.6 or above 1.4 suggests deletion or duplication of the target sequence, respectively. In these positive cases, the MLPA study was extended to the other available members of the family to complete segregation analysis.

To delimit genomic rearrangements found in previous MLPA or WES analyses, aCGH was performed by using the commercial aCGH platform, SurePrint G3 CGH+SNP 2X400k (Agilent Technologies, Inc., Santa Clara, CA, USA), which includes 292,097 CGH probes and 119,091 SNP probes with 3.7 kb of mean space. Results were analyzed by Agilent CytoGenomics software v2.7 using the default analysis method CGH+SNP v2 with the ADM-2 aberration algorithm.

Additionally, to verify whether the duplication region was in tandem, a long PCR strategy was used with the Expand Long Range dNTPack kit (Roche Life Science, Basel, Switzerland), following the manufacturer's instructions. Specific primers for the *PRPF31* exons involved in the breakpoint boundaries were designed (Exon 1F: ACTTGTCCTCATCAGGCTCAG; Exon 6-7R: TCCCAAGGTCACAGTGTCTAG). Polymerase chain reaction amplicons were checked in a 1% agarose gel, using Lambda DNA/*HindIII* as a marker. The amplification of the wild-type allele corresponds with a band of 8.6 kb, while a PCR amplification of 13.7 kb was expected for tandem duplication.

Expression of the *PRPF31* Gene

Total RNA from three individuals carrying *PRPF31* rearrangements and nine wild-type controls were transcribed by using random primers with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) in a final volume of 20 μ L. Quantitative PCR (qPCR) was performed with specific primers (Exon 7-8F: CCAAGATCATGGGTGTGG; Exon 8R: GGCGCAGCACTGAGGTAG) and SYBR Green PCR Master Mix (Roche Life Science) on a LightCycler 480 Real-Time PCR System (Roche Life Science). Quantitative PCR analysis was performed twice and in triplicate. Glyceraldehyde-3-phosphate dehydrogenase was also analyzed in the qPCR experiment as an endogenous gene control. Fluorescence intensity was analyzed by using LightCycler 480 gene scanning software (Roche Life Science), and the $\Delta\Delta C_t$ method was used to calculate relative abundances in *PRPF31* mRNA as fold-change values. The Mann-Whitney test was used to calculate statistical significance between different groups.

MSR1 Copy-Number Analysis

The number of copies of the minisatellite repeat element (MSR1) adjacent to the *PRPF31* promoter was studied in 35 symptomatic and 5 asymptomatic patients from 15 families with causative mutation in the *PRPF31* gene, as previously reported by Rose et al.¹⁵ Additionally, a cohort of 123 control samples was also studied to obtain the minor allele frequency (MAF) of the 4 MSR1 allele. Amplified fragments were

separated in an ABI 3130xl automatic analyzer and analyzed with GeneMapper. Proportion test or Fisher exact test, when the expected frequency was less than 5, was used to calculate statistical significance.

RESULTS

Screening of *PRPF31* Mutations

Our cohort of 211 Spanish patients was sequentially screened for variants in the *PRPF31* gene, following a combination of different methods (Fig. 1). Table 1 summarizes all the mutations reported in this study. Preliminary studies on the implication of *PRPF31* in adRP were initially performed by using SSCP and DGGE in 60 patients, allowing the genetic characterization of two adRP families (RP-0637 and RP-0368; Supplementary Fig. S1), as previously reported.⁹ Further *PRPF31* mutations were identified in 5 of 189 affected adRP families by using the adRP genotyping microarray (RP-2136, RP-0182, RP-0006, RP-1609, and RP-1599; Supplementary Fig. S1). Additionally, comprehensive resequencing of all exons and splicing boundaries of *PRPF31* was performed by means of customized targeted NGS gene panels and WES in 76 and 6 uncharacterized families, respectively,^{18,19,20} allowing the identification of causative variants in six additional families (RP-0361, RP-1541, RP-2496, RP-1970, RP-1318, and RP-0777; Supplementary Fig. S1).

Using this combined approach, 12 different causative mutations in the *PRPF31* gene were found in 13 of 211 unrelated families, yielding a mutation detection rate of 6.2%. We found five splicing, three frameshift, one missense, one small in-frame deletion, one large deletion, and one nonsense mutation in two different families (RP-0006 and RP-1609), two of them being novel (Fig. 2; Table 1).

Analysis of *PRPF31* Rearrangements

In 103 index cases with negative results observed through the NGS custom panel approach, the genetic analysis was extended to perform MLPA, finding *PRPF31* rearrangements in two families (RP-0932 and RP-2426) (Table 1).

First, in family RP-0932, a large pedigree in which RP is segregating in nine individuals from two different branches (Fig. 3), two different types of rearrangements in *PRPF31* were identified (Table 1; Supplementary Fig. S1). Within the same family, a duplication from exons 2 to 5 was found in the index case and her affected father (IV:1 and III:1, respectively), whereas a large deletion from exons 1 to 13 was observed in two affected individuals: her paternal aunt and her cousin (III:12 and IV:12, respectively).

In both rearrangements, the breakpoint regions were defined by using a commercial aCGH of 400K probes that cover the *PRPF31* locus with 34 probes. The deletion size is at least 29.7 kb, encompassing not only *PRPF31* partially but also three contiguous upstream genes (*TFPT*, *NDUFA3*, and part of *OSCAR*), as detailed in Figure 2 and Table 2. The aCGH analysis enabled us to delimit the duplication with a minimum size of 5.1 kb involving only part of *PRPF31* (Fig. 2; Table 2). Additionally, a specific long PCR allowed us to determine that both patients (IV:1 and III:1) carried a partial tandem duplication of *PRPF31* (Supplementary Fig. S2).

To check the behavior of the deletion and the duplication in the *PRPF31* gene—as this is the first time a large duplication has been reported—the expression analysis of this gene was also performed. The *PRPF31* expression was significantly decreased from patients carrying either the deletion (IV:12) or the duplication (IV:1 and III:1), showing levels ranging from

Table 1. Spectrum of *PRPF31* Mutations in adRP Spanish Families

Family	Exon	<i>PRPF31</i> Mutations		References	Technique	Incomplete	
		Nucleotide Change	Protein Change			Penetrance	Segregation
RP-0932 _{del}	Deletion of <i>OSCAR</i> (E1 to E2), <i>NDUFA3</i> , <i>TFPT</i> , and <i>PRPF31</i> (E1 to 13)			This study	MLPA	Yes	Yes
RP-0932 _{dup}	Duplication of <i>PRPF31</i> (E2 to E5)			This study	MLPA	Yes	Yes
RP-2426	Deletion of <i>PRPF31</i> (E1)			This study	MLPA	No	NA
RP-2496	IVS10-1	c.1074-1G>T	Splice	This study	NGS panel	No	NA
RP-0182	IVS11+2	c.1146+2T>A	Splice	This study	adRP chip	Yes	Yes
RP-0777	Deletion of <i>OSCAR</i> (E1 to E2), <i>NDUFA3</i> , <i>TFPT</i> , and <i>PRPF31</i> (E1 to 14)			Almoguera et al., 2015 ²⁰	WES	No	Yes
RP-1318	4	c.322+4_322+7del	Splice	Zhang et al., 2016 ³⁵	NGS panel	Yes	NA
RP-0361	5	c.328_330del	p.Ile110del	de Sousa Dias et al., 2013 ¹⁸	NGS panel	Yes	Yes
RP-2136	IVS6+1	c.527+1G>T	Splice	Chakarova et al., 2006 ³⁶	adRP chip	No	NA
RP-0006	7	c.541G>T	p.Glu181*	Pomares et al., 2010 ³⁷	adRP chip	No	Yes
RP-1609	7	c.541G>T	p.Glu181*	Pomares et al., 2010 ³⁷	adRP chip	Yes	Yes
RP-0637	8	c.770dup	p.Thr258Asp*21	Vithana et al., 2001 ⁴	SSCP/DG-GE	Yes	Yes
RP-0368	8	c.828_829del	p.His276Gln*2	Martinez-Gimeno et al., 2003 ⁹	SSCP/DG-GE	Yes	Yes
RP-1599	9	c.895T>C	p.Cys299Arg	Sullivan et al., 2006 ³⁸	adRP chip	Yes	Yes
RP-1541	9	c.939dup	p.Gly314Arg*10	Fernandez-San Jose et al., 2015 ¹⁹	NGS panel	Yes	Yes
RP-1970	IVS11+2	c.1146+2T>C	Splice	Waseem et al., 2007 ²¹	NGS panel	Yes	Yes

Nucleotide numbering reflects cDNA in the reference sequence NM_015629, according to the guidelines of the Human Genome Variation Society. The initiation codon is codon +1. Different molecular approaches were used to identify pathogenic alleles. References marked in bold show mutations described in our cohort for the first time; NGS panel,^{18,19} adRP chip, autosomal dominant RP genotyping microarray; E, exon; del, deletion; dup, duplication; NA, data not available.

52% to 70% as compared with control samples ($P < 0.001$) (Supplementary Fig. S3).

In family RP-2426, aCGH analysis confirmed a small deletion of a minimum size of 227 bp, which encompasses only exon 1 of the *PRPF31* gene, as previously detected by MLPA (Fig. 2; Table 2).

Breakpoint regions were also defined for family RP-0777 by using the aforementioned aCGH. This family, which was initially negative for adRP genotyping microarray, was finally characterized by WES as carrying a gross deletion encompassing the entire *PRPF31* gene along with other adjacent genes²⁰ (part of the *OSCAR* gene and the entire *NDUFA3* and *TFPT* genes). Thus, we identified a larger deletion of at least 32.2 kb (Fig. 2; Table 2).

Spectrum of *PRPF31* Defects in Our Spanish Population

In a total of 15 families, causative mutations or CNVs in the *PRPF31* gene were found by combining different genetic approaches. In one pedigree, two different gene rearrangements were identified within the same family. Thus, a total of 16 *PRPF31* defects were identified, obtaining a frequency of 7.6% (16/211) in this cohort. The prevalence of coding mutations and genomic rearrangements in the *PRPF31* gene were 5.7% and 1.9%, respectively.

Fifteen different mutations were identified: five splicing mutations (33.3%), three frameshift mutations (20.0%), three large deletions (20.0%), one missense mutation (6.7%), one nonsense mutation (6.7%), one in-frame mutation (6.7%), and one large duplication (6.7%).

Each mutation was found once with the exception of the p.Glu181* mutation, which was recurrent for two unrelated families (RP-0006 and RP-1609). Haplotype analysis using intragenic SNPs and microsatellite markers flanking *PRPF31* was performed in both families. The analysis of the haplotypes

showed a common region sharing a minimum of 1.2 Mb and a maximum of 1.4 Mb, suggesting a founder effect rather than a recurrent mutation (Supplementary Fig. S4).

According to the available data on family history, lack of penetrance was observed in up to 66.7% (10/15) of the families.

Contribution of MSRI Repeat Element to Disease Penetrance of *PRPF31* Mutations

Rose et al.¹⁵ have reported that copy number variation of a cluster of MSRI elements located upstream to the *PRPF31* promoter could have a major genetic role in determining incomplete penetrance of *PRPF31*-associated adRP by acting as modulator of its transcriptional activity. Considering that incomplete penetrance was identified in up to 66.7% of our families, affecting at least 12% of genetically confirmed *PRPF31* carriers, we have evaluated the possible implication of this MSRI repeat element in the disease expressivity in our cohort. Thus, CNV analysis was assessed in 40 *PRPF31* mutation carriers, including 35 symptomatic from 15 different families and 5 asymptomatic individuals from 5 of these families (Supplementary Fig. S1). After genotyping of symptomatic individuals, 54% (19/35) carried the 3/4 MSRI genotype and only 46% (16/35) carried the 3/3 genotype. In this last group, we have also included four hemizygous individuals for the 3 MSRI allele, owing to the presence of large *PRPF31* deletions affecting the MSRI region. No affected patients carried the 4/4 genotype. After considering the high proportion of 4 MSRI allele, the MAF estimation on Spanish population was 0.13. Although the 3/3 (20%), 3/4 (40%), and 4/4 (40%) genotypes were found among the asymptomatic group, the 4 MSRI allele was significantly overrepresented ($P < 0.001$) with respect to ethnically matched controls. These observations were consistent with the hypothesis that the 4-copy allele provides higher *PRPF31* gene expression and could act as a protective factor in

Analysis of the *PRPF31* Gene: A Novel Genomic Rearrangement

IOVS | February 2017 | Vol. 58 | No. 2 | 1049

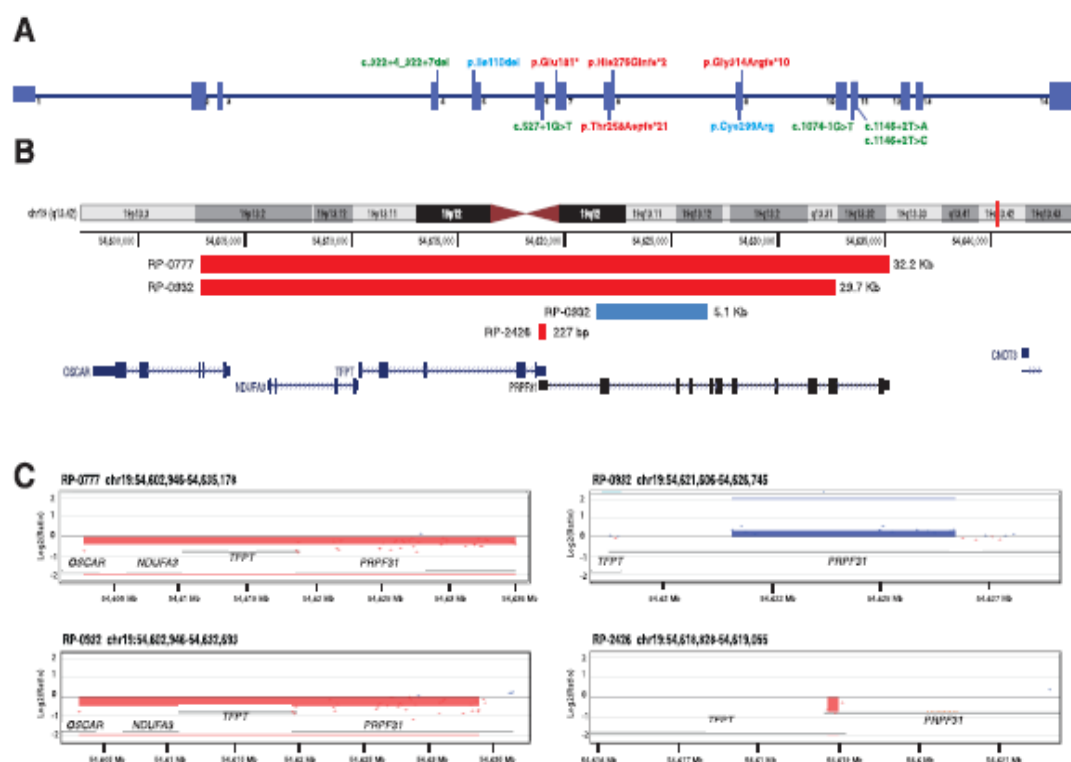


FIGURE 2. Overview of *PRPF31* mutations and large rearrangements identified in our study. (A) Schematic representation of the *PRPF31* gene showing the localization of the coding mutations identified. Exons are indicated by colored rectangles that are wider for the coding regions. Nucleotide numbering reflects cDNA in the reference sequence NM_015629, according to the guidelines of the Human Genome Variation Society. The initiation codon is codon +1. Splice site, missense/in-frame, and nonsense/frameshift are written in green, blue, and red, respectively. (B) Characterization of *PRPF31* rearrangements and other adjacent genes on chromosome 19 in families RP-0777, RP-0932, and RP-2426 by aCGH analysis. Commercial comparative genomic hybridization array was performed to delimit previous MLPA or WBS findings in individuals carrying either deletion or duplication. Red bars represent the genomic positions of the deletions in families RP-0777 (chr19:54602946-54635178), RP-0932, patient IV:12, (chr19:54602946-54632693), and RP-2426 (chr19:54618828-54619055). The blue bar indicates the genomic positions of the duplication (chr19:54621606-54626745) in patient IV:1 in family RP-0932. Schematic representation of the complete intron-exon structure of *PRPF31* and contiguous genes are shown. Exons are indicated by rectangles. (C) Comparative genomic hybridization array data for four individuals are shown. The patient versus reference log2-ratio of each probe is plotted. Dots with log2 ratio approximately -1 indicate a heterozygous deletion (red dots), log2 ratio 0 indicates normal pattern, and +0.6 indicates heterozygous amplification (blue dots). Shaded areas indicate significant gains and losses. Genomic coordinates are based on the Human Genome Assembly hg19.

PRPF31-associated disease expressivity. Lack of penetrance was observed in all families with symptomatic 3/4 *MSR1* alleles (Supplementary Fig. S1).

Genotype-Phenotype Correlation for *PRPF31*

A summary of clinical and genetic data is presented in Table 3 and Supplementary Table S1. Night blindness was the earliest presenting symptom (age, 11.7 ± 6.6 years), followed by visual field constriction (17.2 ± 9.7 years) and visual acuity loss (23.8 ± 13.4 years). Subcapsular cataracts were found in up to 36% of the patients (range: 21–65 years). Apparently, the most severe symptoms were observed in the patient carrying the tandem duplication (IV:1, RP-0932), in terms of onset and course of the disease. However, since we only studied one patient with a duplication, we could not obtain a general genotype-phenotype correlation. Moreover, the cases with missense/in-frame mutations seem to have an earlier onset and rapid evolution to loss of vision. Deletion of flanking genes (*OSCAR*, *NDUFA3*, and *TFPT* genes) was not associated with additional obvious signs.

DISCUSSION

This study provided the prevalence not only for the single-nucleotide variants but also for genomic rearrangements in the *PRPF31* gene in a Spanish adRP cohort. In addition, we described two novel mutations, two new large deletions, and a gross duplication associated with adRP, described for the first time in this study. These results revealed a *PRPF31* mutation prevalence of 7.6%, which is higher than that previously reported in another study on the Spanish population (1.7%).⁹ This prevalence is similar to that found in a French population (6.7%)⁸ and higher than the one observed in the UK,²¹ Japanese,²² and Chinese populations,²³ with frequencies of 5%, 3%, and 1%, respectively. The frequency of genomic rearrangements in the *PRPF31* gene, first reported for the Spanish population, accounts for 1.9%. This prevalence is lower than one previously described²⁴ in a North American population but higher than in France,⁸ where no genomic rearrangements have been found.

Mutational *PRPF31* spectrum in our cohort is very heterogeneous, as we found a total of 16 mutations in 15 different families tested. No hotspot or recurrent mutations

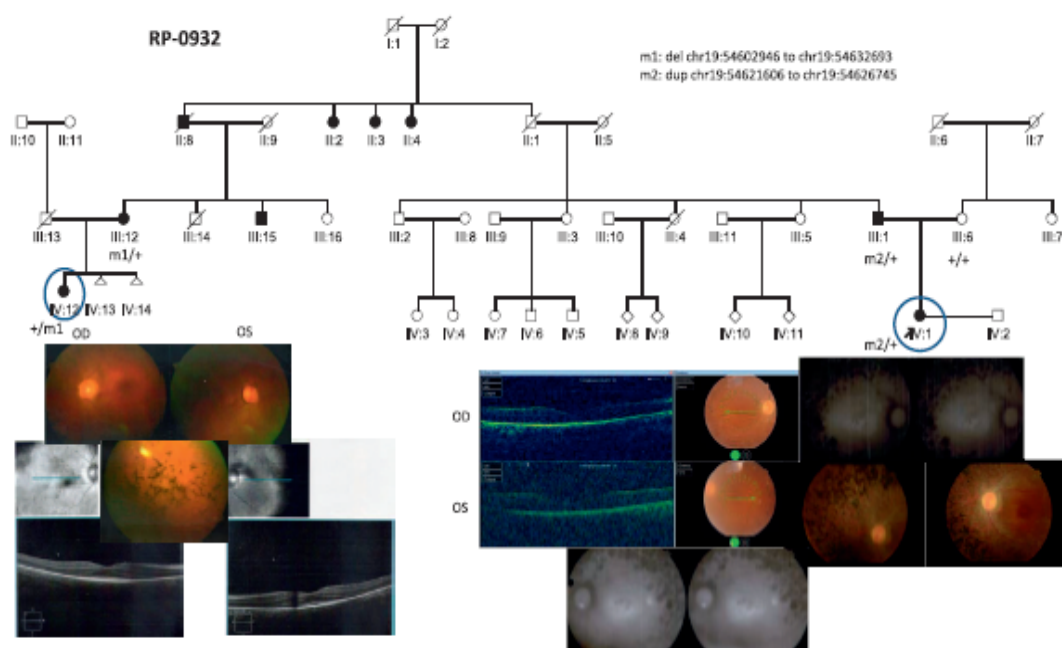


FIGURE 3. Pedigree of family RP-0932 and retinal imaging from the index case and her cousin, who carried duplication and deletion within the *PRPF31* gene, respectively. Ophthalmoscopic and optic coherence tomography (OCT) examination of patients IV:1 (proband, with a duplication within the *PRPF31* gene) and IV:12 (cousin, with a deletion within the *PRPF31* gene). Patient IV:1, at the age of 41 years, showed midperipheral and posterior (outside the arcades) clumps of black pigment in bone spicule configuration, macular alteration, attenuated retinal vessels, atrophy of the retinal pigment epithelium (RPE) with choroidal transparency, and pallor of the optic nerve. Impression of disruption and thinning of the retina (mainly on OS) on OCT. Patient IV:12, at the age of 38 years, showed midperipheral and posterior (outside the arcades) clumps of black pigment in bone spicule configuration, with macular alteration, attenuated retinal vessels, and pallor of the optic nerve. Slight epiretinal membrane on OD. Thinning of retina on OCT. OD, right eye; OS, left eye.

were found, except for the p.Glu181* mutation, a previously known mutation,¹⁹ which was detected in two of our families. Both families have the same geographic origin and share a 1.2-Mb common haplotype, therefore suggesting a founder effect. Additionally, in the RP-0932 family, two different large rearrangements were found in different individuals within the same family.

Most of the mutations found in our population are LOF mutations (e.g., nonsense, splicing, frameshift mutation, or gross rearrangements), presumably causing premature termination codon in alternative transcripts susceptible to being degraded via nonsense-mediated mRNA decay (NMD), as has been previously described by Rio Frio et al.²⁵ For this reason, most of the *PRPF31* mutations seem to cause disease by haploinsufficiency.²⁵ In the RP-0932 family, the *PRPF31* expression was significantly reduced from 30% to 48% as compared to control samples in both deletion and duplication patients. Our finding supports the hypothesis of haploinsuffi-

ciency as the pathogenic mechanism associated with *PRPF31*. In the case of the tandem duplication, the most likely cause of haploinsufficiency is the degradation of aberrant *PRPF31* transcripts through NMD. In all the patients carrying *PRPF31* deletions, the 5'UTR region on the noncoding exon 1, which is predicted to contain several transcriptional regulatory sites, is involved, and as a result, *PRPF31* haploinsufficiency could be explained by dysregulation of gene transcription.

In view of the high genetic heterogeneity and complexity of *PRPF31*-associated mutational mechanisms, molecular studies for adRP require not only a comprehensive resequencing of exons and splicing sites of this gene but also a copy-number assay by means of MLPA or aCGH to detect large rearrangements. MLPA has some advantages including ease of use, high throughput, and relatively small quantities of DNA needed, but on the other hand, aCGH allows a more reliable refining of rearrangements and breakpoints, particularly if other adjacent genes seem to be involved. Although a large deletion in family

TABLE 2. Patients Carrying Large Genomic Rearrangements on 19q13.42, Affecting *PRPF31*

Family	Kind of Rearrangement	Mutation		Size, bp	
		Genomic Coordinates, hg19	Involved Genes	Minimum	Maximum
RP-0777	Deletion	chr19:54,602,946-54,635,178	<i>OSCAR</i> (E1 to E2), <i>NDUFA3</i> , <i>TFPT</i> , and entire <i>PRPF31</i>	32,232	111,184
RP-0932	Deletion	chr19:54,602,946-54,632,693	<i>OSCAR</i> (E1 to E2), <i>NDUFA3</i> , <i>TFPT</i> , and <i>PRPF31</i> (E1 to E13)	29,747	87,435
RP-0932	Duplication	chr19:54,621,606-54,626,745	<i>PRPF31</i> (E2 to E5)	5,139	7,832
RP-2426	Deletion	chr19:54,618,828-54,619,055	<i>PRPF31</i> (E1)	227	18,609

Table 3. Age of Onset of Night Blindness, Visual Field Constriction, Visual Acuity Loss, and Cataract in Different Types of Mutations

Mutation Type	NB (No. of Cases)	VFC (No. of Cases)	VAL (No. of Cases)	% of VAL	Cataract (No. of Cases)	% Cataract
Splice site	12.6 ± 5.0 (5)	17.0 ± 11.4 (3)	25.5 ± 13.4 (2)	67%	38.0 (1)	17%
Missense/in-frame	10.3 ± 6.4 (3)	10.0 ± 4.7 (4)	13.3 ± 7.8 (4)	100%	30.0 (1)	20%
FS/nonsense	10.5 ± 6.1 (7)	21.8 ± 10.1 (6)	37.3 ± 10.3 (4)	75%	44.4 ± 16.9 (5)	63%
Tandem duplication	5.0 (1)	7.0 (1)	20.0 (1)	100%	25.0 (1)	50%
Partial deletion	27.0 (1)	21.0 ± 9.6 (3)	13.0 (1)	33%	50.0 (1)	33%
Total	12.3 ± 6.8	14.9 ± 6.1	20.1 ± 10.0		37.5 ± 10.2	

Data shown as average ± standard deviation. FS, frameshift; NB, night blindness; VAL, visual acuity loss; VFC, visual field constriction.

RP-0777 was identified by WES, in general CNV analyses using NGS data are still constrained owing to technical issues such as coverage and design.

We have successfully applied aCGH for refining *PRPF31* rearrangements that were previously detected by other techniques. Here, we used a commercial array that specifically targets genes and loci previously associated with pathogenic chromosomal aberrations, as in the case of *PRPF31*. Although this design allowed for a high coverage of *PRPF31*, its up- and downstream genomic regions are less covered, so in several cases, we could not establish the exact breakpoints for the largest deletions found in our patients. This was the case of contiguous gene deletions in the RP-0932 and RP-0777 families, for which the deletion seems to start at the same point (Table 2). In the RP-0777 family, the detected deletion was longer than that previously reported,²⁰ including *PRPF31* exon 14 and *OSCAR* exon 2.

Large deletions in *PRPF31* affecting several continuous exons, the whole gene, or even part of *PRPF31* along upstream genes^{24,26} have been described previously, as in our cohort. To our knowledge, this is the first time that two different kinds of *PRPF31* rearrangements—a duplication and a deletion—are reported in different members within the same family and associated with adRP. Interestingly, we also identified a small deletion involving only a single exon of *PRPF31*. In family RP-2426, a deletion of the noncoding exon 1 was detected; thus, a haploinsufficiency mechanism was expected in affected patients due to loss of regulatory regions. Unfortunately, *PRPF31* gene expression studies were not available for this family.

The *PRPF31* gene is located on a region rich in repeat elements, especially Alu repeats. Alu elements are a type of short interspersed element that accounts for 10% of the total genome content, but in chromosome 19 accounts for 26.3%, with Alu being the richest chromosome.²⁷ Alu repeats have been implicated in human genetic diseases by insertion of Alu elements or by nonallelic homologous recombination causing rearrangements. This last mechanism, which is predicted to cause 0.3% of human genetic diseases,²⁸ could be the most likely explanation for the high prevalence of genomic *PRPF31* rearrangements. The mechanism underlying these complex structural rearrangements in family RP-0932 could be nonallelic homologous recombination between nearly identical Alu repeats.

Incomplete penetrance events were found in a large number of families with mutation in the *PRPF31* gene, in particular in 66.7% of the families of our cohort. The mechanism that explains this phenomenon is allelic imbalance with high expression of the wild-type allele, compensating for the mutant allele in asymptomatic carriers.^{29,30} A possible role of *PRPF31* modifier genetic factors, such as *CNOT3* or the MSRI repeat element upstream of *PRPF31*,¹² have been hypothesized to explain this overexpression event.^{13–15} These modifier genetic factors are involved in transcriptional

repression of the *PRPF31* gene.^{13–15} However, the exact mechanism is still unclear.

The 3- and 4-copies of the MSRI minisatellite showed strong differences in the modulation of the *PRPF31* gene transcription, which could be associated with incomplete penetrance in *PRPF31*-associated RP.¹⁵ Consistent with the hypothesis that the 4 MSRI allele seems to be a major protective factor for disease expressivity,¹⁵ asymptomatic individuals from our cohort had a frequency for 4 MSRI allele significantly higher than that of the control population. Rose et al.¹⁵ have reported that all symptomatic patients in their study are homozygous for the 3 MSRI allele, predicted to have a decreased *PRPF31* expression, and thus developing the disease. However, 33% of our families (RP-0932, RP-0368, RP-0637, RP-0361, and RP-1599) had at least one symptomatic individual that was heterozygous 3/4; this fact is contradictory with that seen by Rose et al.,¹⁵ reflecting a higher frequency of 4 MSRI allele than previously reported. In these heterozygous symptomatic cases, it is expected that the *PRPF31* gene expression will be reduced and then symptoms will develop if the wild-type allele is in *cis* with the 3-repeat MSRI allele and the 4 MSRI allele is in *cis* with the mutant allele, so the 4-copy does not seem to act as a protective factor. Conversely, in heterozygous asymptomatic individuals, the protective 4-allele should be in *cis* with the wild-type allele in order to maintain a high expression of *PRPF31*, thus compensating the mutated allele. Unfortunately, it was not possible to fully confirm this hypothesis in our study, given that the complete phasing of alleles could not be inferred in most families owing to lack of familial information and/or sample availability. However, we observed two clear exceptions for this hypothesis in two heterozygous 3/4 asymptomatic individuals, the III:6 and III:9 individuals from families RP-0361 and RP-1599, respectively, where the 4 MSRI allele is necessarily in *cis* with the mutant allele as the haplotype analysis showed (Supplementary Fig. S1). In these two cases, the MSRI cannot explain the asymptomatic status, so something else must be present outside the 1.9 Mb region around the *PRPF31* gene to explain the incomplete penetrance. This last observation is in agreement with a previous hypothesis of a polygenic control of *PRPF31* transcriptional regulation in the incomplete penetrance of *PRPF31*-associated RP.

An ophthalmic review of 26 patients from 15 different families carrying *PRPF31* mutations was performed, representing the largest phenotypic description for this gene to our knowledge. The clinical signs observed in family RP-0932 seem to be more severe in the patients carrying the duplication than in the patients carrying the deletion in *PRPF31*. Thus, this family represents a good example for phenotypic variability associated with different *PRPF31* defects. In our cohort, patients carrying a missense mutation, in contrast to those patients carrying a nonsense mutation or a frameshift mutation, did not necessarily have a milder phenotype as previously reported by Xu et al.³¹ Though these results are preliminary and cannot be used to establish a clear genotype-

phenotype correlation because clinical information for patients carrying *PRPF31* defects was limited in some cases, it seems that the duplication and the missense mutations could produce a more severe phenotype than the other kind of mutations. Patients from families RP-0932 and RP-0777 with deletion in additional upstream genes did not show additional clinical findings apart from RP, which is consistent with previous observations.³²⁻³⁴

In conclusion, in light of the high genetic heterogeneity of *PRPF31* mutations, it is necessary to screen all exons and splicing boundaries and also perform copy-number assays, as large duplications or deletions in the *PRPF31* gene can be also identified. The identification of novel pathogenic mechanisms underlying *PRPF31*-related RP and genotype-phenotype correlation is important not only for current diagnostics but also for genetic counseling. In addition, it is particularly important for the development of new gene-dependent therapies and further inclusion of the patients in human gene-specific clinical trials.

Acknowledgments

The authors thank the Genetics and Ophthalmology Departments of the Fundación Jiménez Díaz University Hospital (Madrid) and especially Asunción Giménez, Cristina Nieto, Cristina Villaverde, and Olga Zurita for their technical assistance. They also thank Oliver Shaw for English review and editing the manuscript. They thank all patients and doctors who participated in the study, especially María Asunción López Ariztegui from Hospital Universitario Cruces, Bilbao, Guimar Pérez de Nancrales from Hospital Universitario Araba-Txagorritxu, Vitoria, and Cristina Irigoyen Laborra from Hospital Universitario de Donostia, San Sebastián.

Supported by several grants from the Instituto de Salud Carlos III (ISCIII) from the Spanish Ministry of Health, including CIBERER (06/07/0036), HS - FEDER (European Regional Development Fund) (PI013/00226), and IIS-FJD Biobank PT13/0010/0012. In addition, the Spanish National Organization of the Blind (ONCE) and the Spanish Fighting Blindness Foundation (FUNDALUCE) also supported our work. IMM and RS-A are sponsored by the Rio Hortega and Sara Borrell Postdoctoral Programs (CD14-00079 and CD12-00676) from ISCIII/FEDER. MCo is supported by the Miguel Servet Program (CP12/03256) from ISCIII. RP-C is supported by Fundación Conchita Rabago (FCR), and LRJdS was supported by the CAPES Foundation, Ministry of Education of Brazil.

Disclosure: I. Martín-Merida, None; R. Sánchez-Alcudia, None; P. Fernández-San Jose, None; F. Blanco-Kelly, None; R. Pérez-Carro, None; L. Rodríguez-Jacó da Silva, None; B. Almoguera, None; B. García-Sandoval, None; M.I. López-Molina, None; A. Avila-Fernandez, None; M. Carballo, None; M. Corton, None; C. Ayuso, None

References

- Hamel C. Retinitis pigmentosa. *Orphanet J Rare Dis*. 2006;1:40.
- Ayuso C, Milan JM. Retinitis pigmentosa and allied conditions today: a paradigm of translational research. *Genome Med*. 2010;2:34.
- Ayuso C, García-Sandoval B, Najera C, Valverde D, Carballo M, Antinolo G. Retinitis pigmentosa in Spain: The Spanish Multicentric and Multidisciplinary Group for Research into Retinitis Pigmentosa. *Clin Genet*. 1995;48:120-122.
- Vithana EN, Abu-Safieh I, Allen MJ, et al. A human homolog of yeast pre-mRNA splicing factor, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell*. 2001;8:375-381.
- Makarova OV, Makarov EM, Liu S, Vornlocher HP, Lührmann R. Protein 61K, encoded by a gene (PRPF31) linked to autosomal

- dominant retinitis pigmentosa, is required for U4/U6U5 tri-snRNP formation and pre-mRNA splicing. *EMBO J*. 2002;21:1148-1157.
- Liu S, Li P, Dybkov O, et al. Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science*. 2007;316:115-120.
- Cao H, Wu J, Lam S, et al. Temporal and tissue specific regulation of RP-associated splicing factor genes PRPF3, PRPF31 and PRPF8—implications in the pathogenesis of RP. *PLoS One*. 2011;6:e15860.
- Audo I, Bujakowska K, Mohand-Said S, et al. Prevalence and novelty of PRPF31 mutations in French autosomal dominant rod-cone dystrophy patients and a review of published reports. *BMC Med Genet*. 2010;11:145.
- Martínez-Gimeno M, Gamundi MJ, Hernán I, et al. Mutations in the pre-mRNA splicing-factor genes PRPF3, PRPF8, and PRPF31 in Spanish families with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2003;44:2171-2177.
- Stenson PD, Ball EV, Mort M, et al. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat*. 2003;21:577-581.
- McGee TL, Devoto M, Ott J, Berson EL, Dryja TP. Evidence that the penetrance of mutations at the RP11 locus causing dominant retinitis pigmentosa is influenced by a gene linked to the homologous RP11 allele. *Am J Hum Genet*. 1997;61:1059-1066.
- Rose AM, Bhattacharya SS. Variant haploinsufficiency and phenotypic non-penetrance in PRPF31-associated retinitis pigmentosa. *Clin Genet*. 2016;90:118-126.
- Venturini G, Rose AM, Shah AZ, Bhattacharya SS, Rivolta C. CNOT3 is a modifier of PRPF31 mutations in retinitis pigmentosa with incomplete penetrance. *PLoS Genet*. 2012;8:e1003040.
- Rose AM, Shah AZ, Venturini G, Rivolta C, Rose GE, Bhattacharya SS. Dominant PRPF31 mutations are hypostatic to a recessive CNOT3 polymorphism in retinitis pigmentosa: a novel phenomenon of "linked transacting epistasis." *Ann Hum Genet*. 2014;78:62-71.
- Rose AM, Shah AZ, Venturini G, et al. Transcriptional regulation of PRPF31 gene expression by MSR1 repeat elements causes incomplete penetrance in retinitis pigmentosa. *Sci Rep*. 2016;6:19450.
- Fernández-SanJose P, Blanco-Kelly F, Corton M, et al. Prevalence of rhodopsin mutations in autosomal dominant retinitis pigmentosa in Spain: clinical and analytical review in 200 families. *Acta Ophthalmol*. 2015;93:e38-e44.
- Blanco-Kelly F, García-Hoyos M, Corton M, et al. Genotyping microarray: mutation screening in Spanish families with autosomal dominant retinitis pigmentosa. *Mol Vis*. 2012;18:1478-1483.
- de SousaDias M, Hernán I, Pascual B, et al. Detection of novel mutations that cause autosomal dominant retinitis pigmentosa in candidate genes by long-range PCR amplification and next-generation sequencing. *Mol Vis*. 2013;19:654-664.
- Fernández-SanJose P, Corton M, Blanco-Kelly F, et al. Targeted next-generation sequencing improves the diagnosis of autosomal dominant retinitis pigmentosa in Spanish patients. *Invest Ophthalmol Vis Sci*. 2015;56:2173-2182.
- Almoguera B, Li J, Fernández-SanJose P, et al. Application of whole exome sequencing in six families with an initial diagnosis of autosomal dominant retinitis pigmentosa: lessons learned. *PLoS One*. 2015;10:e0133624.
- Waseem NH, Vaclavik V, Webster A, Jenkins SA, Bird AC, Bhattacharya SS. Mutations in the gene coding for the pre-mRNA splicing factor, PRPF31, in patients with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2007;48:1330-1334.

22. Sato H, Wada Y, Itabashi T, Nakamura M, Kawamura M, Tanai M. Mutations in the pre-mRNA splicing gene, *PRPF31*, in Japanese families with autosomal dominant retinitis pigmentosa. *Am J Ophthalmol*. 2005;140:537-540.
23. Lim KP, Yip SP, Cheung SC, Leung KW, Lam ST, To CH. Novel *PRPF31* and *PRPH2* mutations and co-occurrence of *PRPF31* and *RHO* mutations in Chinese patients with retinitis pigmentosa. *Arch Ophthalmol*. 2009;127:784-790.
24. Sullivan LS, Bowne SJ, Seaman CR, et al. Genomic rearrangements of the *PRPF31* gene account for 2.5% of autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2006;47:4579-4588.
25. Rio Frio T, Wade NM, Ransijn A, Berson EL, Beckmann JS, Rivolta C. Premature termination codons in *PRPF31* cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. *J Clin Invest*. 2008;118:1519-1531.
26. Eisenberger T, Neuhaus C, Khan AO, et al. Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: the example of retinal dystrophies. *PLoS One*. 2013;8:e78496.
27. Grover D, Mukerji M, Bhatnagar P, Kannan K, Brahmachari SK. Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. *Bioinformatics*. 2004;20:813-817.
28. Deininger PL, Batzer MA. Alu repeats and human disease. *Mol Genet Metab*. 1999;67:183-193.
29. Vithana EN, AbuSafieh I, Pelosini L, et al. Expression of *PRPF31* mRNA in patients with autosomal dominant retinitis pigmentosa: a molecular clue for incomplete penetrance? *Invest Ophthalmol Vis Sci*. 2003;44:4204-4209.
30. Rivolta C, McGee TL, Rio Frio T, Jensen RV, Berson EL, Dryja TP. Variation in retinitis pigmentosa-11 (*PRPF31* or *RP11*) gene expression between symptomatic and asymptomatic patients with dominant *RP11* mutations. *Hum Mutat*. 2006;27:644-653.
31. Xu F, Sui R, Liang X, Li H, Jiang R, Dong F. Novel *PRPF31* mutations associated with Chinese autosomal dominant retinitis pigmentosa patients. *Mol Vis*. 2012;18:3021-3028.
32. AbuSafieh I, Vithana EN, Mantel I, et al. A large deletion in the *adRP* gene *PRPF31*: evidence that haploinsufficiency is the cause of disease. *Mol Vis*. 2006;12:384-388.
33. Rose AM, Mukhopadhyay R, Webster AR, Bhattacharya SS, Waseem NH. A 112 kb deletion in chromosome 19q13.42 leads to retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2011;52:6597-6603.
34. Köhn L, Bowne SJ, Sullivan LS, et al. Breakpoint characterization of a novel approximately 59 kb genomic deletion on 19q13.42 in autosomal-dominant retinitis pigmentosa with incomplete penetrance. *Eur J Hum Genet*. 2009;17:651-655.
35. Zhang Q, Xu M, Verriotto JD, et al. Next-generation sequencing-based molecular diagnosis of 35 Hispanic retinitis pigmentosa probands. *Sci Rep*. 2016;6:2792.
36. Chakarova CE, Cheminkova S, Tournev I, et al. Molecular genetics of retinitis pigmentosa in two Romani (Gypsy) families. *Mol Vis*. 2006;12:909-914.
37. Pomares E, Riera M, Permanyer J, et al. Comprehensive SNP-chip for retinitis pigmentosa-Leber congenital amaurosis diagnosis: new mutations and detection of mutational founder effects. *Eur J Hum Genet*. 2010;18:118-124.
38. Sullivan LS, Bowne SJ, Birch DG, et al. Prevalence of disease-causing mutations in families with autosomal dominant retinitis pigmentosa: a screen of known genes in 200 families. *Invest Ophthalmol Vis Sci*. 2006;47:3052-3064.

B. Towards the mutational landscape of autosomal dominant retinitis pigmentosa: a comprehensive analysis of 258 Spanish families. Artículo en revisión en la revista IOVS.